PRIMORDIAL PERTURBATIONS IN THE UNIVERSE: THEORY, DETECTION, AND IMPLICATION FOR INFLATION

by

Liang Dai

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

August, 2015

Abstract

Observation supports the theory that the large-scale structure in the Universe was seeded by primordial curvature perturbations generated during inflation. Measuring the departure from gaussian statistics in the primordial curvature perturbation may reveal the underlying physical mechanism of inflation. Furthermore, a primordial background of gravitational waves can also be produced during inflation, and their imprints may be observed in the recent Universe. Current and upcoming cosmological surveys call for improved theoretical modelling of these primordial perturbations in the hope of detecting new physics. The Total Angular Momentum formalism provides a unified analytic framework to compute for scalar, vector and tensor perturbation fields with manifest symmetry under rotation of the sky, facilitating the calculation of spherical harmonic observables in all-sky surveys. New dynamical degree of freedom during inflation may correlate with the curvature perturbation in a non-trivial scalar-scalar-fossil bispectrum, resulting in a departure from statistical isotropy in the fluctuations of the cosmic microwave background. Large-scale primordial gravita-

tional waves cause polarization rotation of the cosmic microwave background photons. They also modulate a quadrupolar direction dependence in the local clustering power spectrum of small-scale matter perturbation modes. This phenomenon may provide new observational avenue to search for those relic waves. The formalism of Conformal Fermi Coordinates provides a general relativistic framework to study the local impact of long-wavelength perturbations on the dynamics of short-wavelength modes. Applying it to long-wavelength scalar perturbation, we find that general relativity predicts no scale-dependent bias of galaxy clustering in single-field inflation.

Acknowledgments

Foremost, I am sincerely grateful to my academic advisor Prof. Marc Kamionkowski. He is a physicist with immense knowledge, deep insight, and persistent passion. I am thankful for his guidance into the study of cosmology — the farthest reach of human intellectual endeavor through space and time — during the golden age of the discipline. I am also thankful for his patience and and support over the years during my research. This dissertation could not have been written without him. I am deeply indebted to his scholarly cultivation without which I could not have been prepared to pursue the career of a professional scientist.

I would also like to express my gratitude to the reminder of my dissertation committee: Prof. Jared Kaplan, Prof. Charles Bennett, Prof. Darrell Strobel, and Prof. David Yarkony, for their insightful questions and comments during the preparation and presentation of my dissertation. Those have inspired me to think deeper and broader about my research.

Moreover, I am thankful to Prof. Kirill Melnikov, Prof. David Kaplan, Prof.

ACKNOWLEDGMENTS

Joseph Silk, Prof. Nima Arkani-hamed, and many others, who happily invited me to join interesting research projects and had insightful discussions on physics with me. I would also like to thank my postdoctoral colleagues and collaborators, from a long but not exhaustive list: Donghui Jeong, Jens Chluba, Joseph Pradler, Junpu Wang, Fabian Schmidt, Enrico Pajer, Ely Kovetz, Alvise Racanelli, Yacine Ali-Haïmoud, Mark Neyrinck, Daniel Grin, Xin Wang, Amanda Yoho, Fabrizio Caola, and many others. They offered me a lot of help and guidance in addition to my advisor.

I have spent the memorable years at Johns Hopkins University with my school companions: Arpit Gupta, Thomas Zorawski, Jingsheng Li, Yaofu Zhou, Cyrus Faroughy, Lin Yang, Zhilei Xu, Matthew Walters, Christopher Brust, Yuan Wan, Patrick Breysse, Julian Muñoz, Nikhil Anand, Ting-Wen Lan, and many others. We shared with each other ups and downs in our study and life. Words are not enough to express my genuine gratitude for their generous help and encouragement.

I owe a great deal to my parents for their constant understanding and selfless support from the other side of the Earth. They always make me feel not lonely whenever I am away from home and alone.

Last but not the least, I would like to thank my girlfriend Jingnan for her love and care. Meeting with her at Johns Hopkins adds to the best of my memory throughout the years I spent in Baltimore.

Dedication

For my mom and dad.

Contents

List of Figures

Chapter 1

Introduction

The physical Universe appears lumpy. Modern astronomical surveys of galaxies, galaxy clusters, quasars, gas clouds and gravitational lenses (e.g. [1–4]) all have revealed that across cosmological length scales $\gtrsim 100$ Mpc, luminous and dark matter are not uniformly distributed, but cluster together in clumps. Furthermore, the relic radiation background of the hot Big Bang coming from when the Universe was only roughly 300,000 years old, as powerful microwave telescopes have uncovered (e.g. [5–8]), also exhibits fine anisotropies in its temperature and polarization. An abundance of cosmological measurements have established the consensus that our Universe is an almost homogeneously and isotropically expanding spacetime with flat geometry, but is weakly perturbed by density fluctuations since the beginning.

The imperfect Universe is more intriguing to physicists, as it is believed that the cosmological perturbations hold secrets to the very infancy of the Universe, when it was much

smaller than it is at present and had a much higher energy density than what can be created in our greatest laboratories today. Over the decades, a leading theory has emerged that the very early Universe underwent inflation, an epoch of nearly exponential expansion of space.[9–12] This paradigm provides a natural explanation for the acausal initial condtion and flat geometry of the Universe. During the inflationary epoch, quantum field theory in curved spacetime further makes the remarkable prediction that vacuum fluctuations in the energy density are unstable against the quasi-exponential expansion.[13–16] Normal modes of fluctuation with different comoving wavelengths were quickly stretched to scales larger than the Hubble scale during inflation, and were amplified to a level of one part in a hundred thousand, providing an initial condition for the evolution of the post-inflationary Universe with superhorizon correlations.

From a theoretical point of view, minimal models with fewest possible content of new physics are favorable. They are the *single-field slow-roll* inflation models. In the classical picture, a scalar inflaton field ϕ is the only major physical degree of freedom, and it has a homogeneous solution $\phi(\mathbf{x}, t) = \phi(t)$ that rolls down along a shallow slope of the potential $V(\phi)$ in the field space, which drives a quasi-exponential expansion. Then inhomogeneous fluctuations in the inflaton field seed the density perturbations. Those models are also referred to as *standard single-clock* inflation models,[17] for the reason that the perturbations are generated such that all horizon patches have identical post-inflationary expansion history, but are simply mis-synchronized with each other, i.e. the perturbations are adiabatic. This can be expressed in terms of an end result for the spacetime metric toward the end of

inflation,

$$ds^2 \;=\; -dt^2 + a^2(t)\exp\left[2\zeta(\mathbf{x})\right]\delta_{ij}dx^i dx^j, \tag{1.1}$$

where $\zeta(\mathbf{x})$ is the primordial *curvature perturbation*, a stochastic variable quantifying the mis-synchronization in e-folds of expansion. If one performs a Fourier transformation $\zeta(\mathbf{x}) \sim \int d^3\mathbf{k}\,\exp[i\mathbf{k}\cdot\mathbf{x}]\,\zeta(\mathbf{k})$, then single-field slow-roll models make the generic prediction that to leading order $\zeta(\mathbf{k})$ is a Gaussian random field obeying a nearly scale-invariant power spectrum

$$\left\langle \zeta(\mathbf{k})\zeta(\mathbf{k}')\right\rangle \;=\; (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')\,\frac{2\pi^2\Delta_\zeta^2}{k^3}\left(\frac{k}{k_0}\right)^{n_s-1}, \tag{1.2}$$

where $\Delta_\zeta^2 \sim 10^{-10}$ gives the mean squared fluctuation, and the spectral index $n_s - 1 \sim 0.04$ quantifies the departure from exact exponential inflation.[6] Besides, k_0 is some pivot scale of experimental choice. Parametrizing our ignorance of the inflationary physics, Δ_ζ^2 and n_s are among the six cosmological parameters of the concordance model, and are the most crucial to the physics of cosmological perturbations. The other parameters, including the baryon density $\Omega_m h^2$, the cold dark matter density $\Omega_c h^2$, the present Hubble constant H_0, and the optical depth of cosmic reionization $\tau_{\rm re}$, are independent model input. So far, the six-parameter concordance model with the inflationary mechanism seems to stand all observational tests.

This theory successfully explains the large-scale structure observed in the late-time Universe. The primordial curvature perturbation $\zeta(\mathbf{x})$ is equivalent to a primordial gravitational

potential field (often referred to as the scalar metric perturbation) $\Phi(\mathbf{x}) = -(3/5)\zeta(\mathbf{x})$ when the conformal-Newtonian coordinate system is used. The mis-synchronization of expansion then also results in inhomogeneous density for radiation, baryons, dark matter and so on in that frame. Since the cosmic expansion decelerates in the late-time Universe, the perturbation wavelength eventually becomes shorter than the Hubble length. When that happens, potential and density perturbations start to evolve under the influence of gravitational attraction, pressure and other forces, following a set of linearized gravitational equations and transport equations. For example, the radiation fluctuation at the time of cosmic recombination is observed to be the cosmic microwave background (CMB) anisotropies on the sky, which provides a fantastic probe to accurately measure the statistics of primordial perturbation.

Furthermore, the post-recombination Universe is dominated by cold dark matter, which behaves as a non-relativistic, pressure-free fluid. Due to instability under gravitational attraction, fractional fluctuation in the dark matter density gradually grows over time and eventually becomes order unity, so that dark matter collapses into dense, bound objects called halos, which harbor luminous galaxies that form after a series of complex astrophysical processes. Therefore, the spatial distribution of various astrophysical objects also encodes a great wealth of information about the primordial curvature perturbation.

Of course, the simpleness of single-clock inflation scenarios do not necessarily imply that they are the reality. Despite the spectacular triumph of the inflation theory, detailed physics underlying the inflationary mechanism is poorly understood or constrained. To the best of our knowledge, there are a large number of viable models on the market,[18,19] with severe

observational degeneracy. Still, it can be extremely rewarding to continue the investigation and gain further insight into the inflationary physics. During the inflationary epoch, huge energy scales beyond the reach of terrestial experiments were involved, possibly related to new dynamical degrees of freedom unknown to particle physicists. It is even more exciting that a correct understanding of the inflationary physics is likely to require a combination of Quantum Mechanics and General Relativity, a fundamental question pursued for decades in theoretical physics.

A promising probe to distinguish between different inflation models is departure from Gaussian statistics in the curvature perturbation. The prediction of Gaussian initial perturbation only applies under the approximation that the relevant degree of freedom is a free quantum field, although in reality self-interaction or interaction with other fields are ubiquitous in the action. The simplest observable of primordial non-Gaussianity is a bispectrum, or the Fourier transform of the curvature three-point correlation function $\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle$. It turns out that in single-field slow-roll models, bispectrum non-Gaussianity of the curvature perturbation is only at the level of 10^{-7},[20] which is too small to have any noticable impact in the late-time Universe. However, alternative theories to single-field slow-roll models allow for the possibility of primordial non-Gaussianity at a detectable level. In those models, the scale and angle dependence of the bispectrum can encode information about the mass and the spin of any new degree of freedom, and hence can be used to distinguish between models.[21]

Phenomenologically, primordial non-Gausssianity contributes to bispectra in the CMB anisotropies, in galaxy clustering, in cosmic 21cm fluctuations and so on, with distinct bis-

pectrum shapes. Further, it can generate a scale-dependent clustering bias for galaxies that is enhanced on large scales,[22] or a similar enhancement in the intrinsic shape alignment of galaxies.[23] Thanks to the rapid development of observational techniques, we are beginning to witness the onset of a golden age in which huge amount of observational data of high quality will be acquired from current or incoming surveys. Those ambitious observational projects include *Euclid*,[24] *WFIRST*,[25] *HETDEX*,[26] *SKA*,[27] *LSST*,[28] and *SPHEREX*[29] and many others. Given the observational prospect, improved theoretical understanding of the cosmological perturbations are in need for successful interpretation of the data.

In addition to the primordial curvature perturbation, which seeds density fluctuation, inflation has another generic prediction. If spacetime fluctuation itself can be quantized, like any other fundamental dynamical system in nature, gravitational waves γ_{ij} are also spontaneously produced in quasi-de Sitter spacetime as the normal modes are stretched to superhorizon scales.[30–33] Those are the spin-two transverse, trace-free excitations of the spatial metric,

$$ds^2 = -dt^2 + a^2(t)\left[\delta_{ij} + \gamma_{ij}\right] dx^i dx^j, \tag{1.3}$$

the unique propagating degrees of freedom of spacetime geometry. In cosmology, gravitational waves are also referred to as the tensor metric perturbation. Analogous to the curvature perturbation ζ, inflation predicts a nearly scale-invariant gravitational wave power

spectrum per polarization state,

$$\left\langle \gamma_s(\mathbf{k})\gamma_{s'}(\mathbf{k}')\right\rangle \;=\; (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')\,\delta_{ss'}\,r\,\frac{2\pi^2\Delta_\zeta^2}{k^3}\left(\frac{k}{k_0}\right)^{n_t}, \tag{1.4}$$

where the tensor spectral index n_t can differ from the scalar spectral index $n_s - 1$. The tensor-to-scalar ratio $r < 0.12$[34] gives the mean squared amplitude relative to that of the curvature perturbation, which is determined by the Hubble scale during inflation.

If a nearly scale-invariant spectrum of primordial gravitational waves from inflation is detected, it will not only be a further verification of General Relativity, but will also be a strong piece of evidence supporting the inflation paradigm. Although in linear perturbation theory, a primordial gravitational wave background has no effect on the formation of large-scale structure of the Universe, they are expected to induce temperature and polarization anisotropies in the CMB. The claimed detection of degree-scale B-mode polarization anisotropies induced by inflationary gravitational waves in the CMB by *BICEP2/Keck Array*[35] was later on questioned for a probable contamination from the polarized dust emission in our own Milky Way.[34] Still, upcoming experiments including *CLASS*,[36] *POLARBEAR*,[37] *BICEP3*,[38] *EBEX*[39] and *PIXIE*,[40] with unprecedented sensitivity and better foreground removal will continue to search for signatures of inflationary gravitational waves in the CMB. Depending on the amplitude, solar-scale gravitational waves in the inflationary background may also be probed in the future using space-based laser interferometry.[41]

Given the tremendous scientific interest in using cosmological perturbations as a proxy to probe the physics of very early Universe, this Dissertation will be mainly focusing on new

developments in the theory and in the phenomenology of primordial perturbations. It will be a compilation of five original contributions to the subject, each discussed in a separate Chapter. Since each Chapter will be prepared as a self-contained, separate presentation of a particular topic as much as possible, more specific introductory remarks will be given separately at the beginning of each Chapter. Still, the topics of the five Chapters are to some extent related to one another. Therefore, the order in which the five Chapters are presented is mainly based on the consideration of prerequisite knowledge.

We now outline the organization of this Dissertation. In Chapter 2 we develop the Total Angular Momentum (TAM) formalism, for perturbation fields of scalar, vector and tensor type. Essentially a mathematical formalism of spherical wave expansion, it is an alternative to the Fourier expansion formalism traditionally adopted in computations of cosmological perturbations. Especially suitable for analysis of all-sky cosmological observables, such as the temperature of the CMB, and the angular number density of galaxies, the TAM formalism has the advantage that rotational symmetry with respect to the central observer is always manifest. Making contact with spherical harmonic observables in all-sky surveys therefore becomes straightforward. The formalism gives up the translational invariance of the physics, for practical consideration, because in large-scale cosmological surveys we may only observe physics along the past light cone, which is subject to significant evolution with redshift. We discover an algebraic method using irreducible operators to relate the TAM construction for vector and tensor fields to that for scalar fields. While the former is mathematically more sophisticated, the latter is the familiar Fourier-Bessel expansion. Further, in the formalism we implement the transverse-longitudinal decomposition for vector

fields, and scalar-vector-tensor decomposition for tensor fields. This Chapter will contain a comprehensive collection of useful, compact mathematical results for future reference.

In Chapter 3 we discuss the imprint in the initial condition from new dynamical degrees of freedom possibly present during inflation. Generic interaction between the new field and the curvature perturbation field can lead to a nonzero curvature-curvature-new–field bispectrum, which appears as a modulation of the curvature two-point correlation by an invisible degree of freedom. The stochastic nature of the new field, furthermore, would imply an effective four-point correlation function for the curvature perturbation, beyond the contribution from Gaussian statistics. This is referred to as the fossil effect, because the existence of the new field can be sought for through possible distortion to the primordial curvature two-point statistics, even if the new field is dynamically decoupled from the other elementary particles during the late-time Universe. We study the observational capability of detecting departure from statistical isotropy in the CMB anisotropies due to the fossil bispectrum. Using the formalism of bipolar spherical harmonics, we construct optimized estimators for bipolar spherical harmonics, and predict the detecting sensitivity for cosmic-variance-limited CMB experiments such as *Planck*. In particular, we highlight the utility of parity-odd bipolar spherical harmonics as unique diagnostic for the spin quantum number of the fossil field.

In Chapter 4 we discuss gravitational lensing of the CMB temperature and polarization by a background of inflationary gravitational waves on large scales, which is analogous to gravitational lensing by foreground matter inhomogeneities. Using linear perturbation theory, we first present a general discussion on the impact of general metric perturbation on

photon propagation, including the effects on energy, on momentum, and on the polarization state. Utilizing the TAM formalism, we then derive harmonic-space lensing kernels to quantify the distortions to CMB temperature and polarization angular power spectra due to general weak gravitational lensing. Our new contribution to the literature is accounting for the gravitational Faraday rotation of the polarization state in the presence of vector or tensor metric perturbations. We demonstrate that this rotation is fully coherent with the parity-odd deflection of the line of sights, and therefore must be supplemented to the previous results on lensed CMB polarization. This leads to corrected formula for the lensed polarization angular power spectra. In the case of lensing by a background of primordial gravitational waves, we show that ignoring the effect gravitational Faraday rotation results in an overestimation of the secondary B-mode polarization pattern converted from primary E-modes.

Chapter 5 has focus on the imprints of large-scale primordial gravitational waves in matter clustering during the epoch of large-scale structure formation. In standard single-field inflation, the correlation between a long-wavelength gravitational wave and two short-wavelength scalar modes before horizon entry is constrained by the tensor-scalar-scalar consistency relation, which requires that a local observer measures the same scalar two-point correlation as in the absence of the long-wavelength gravitational wave, when compared at the same physical scale rather than coordinate scale. However, we show that after horizon entry the gravitational wave oscillates, and the growth of matter perturbations on short scales is influenced by a time-dependent tidal force. This eventually results in the following fossil imprint: the locally measured two-point correlation of the matter density field (or

galaxy number density) on small scales has a quadrupolar dependence on direction, which is modulated by the gravitational-wave amplitude over large scales. The fractional strength of the quadrupole is equal to the superhorizon amplitude of the gravitational wave. We demonstrate that this fossil effect is permanently imprinted in the large-scale structure even after the primordial gravitational wave redshifts away as the Universe continues to expand. Based on this effect, it is suggested that with massive galaxy surveys in the future, large-scale intrinsic alignment of galaxies may become a promising probe of primordial gravitational waves.

Finally, in Chapter 6 we study how a long-wavelength perturbation in the Universe can locally impact the short-scale dynamics. We show that in the vicinity of given free-falling observer, the Conformal Fermi Coordinates (CFC) can be constructed, which resembles a homogeneously expanding Friedmann-Lemaitre-Robertson-Walker (FLRW) universe. The validity of this "separte universe" picture is limited to scales larger than the sound horizon of the cosmic fluid, and relies on the absence of anisotropic stress. We find that the isotropic effect of a long-wavelength scalar/density perturbation is a modification of the expansion rate and an effective spatial curvature, and the anisotropic effect is completely captured by a tidal potential in the Newtonian form without explicit general relativistic correction, as long as the potential is given in the conformal Newtonian gauge. We then apply the CFC formalism to compute how the growth of subhorizon-scale matter perturbation is modulated by a large-scale density perturbation of arbitrary wavelength, a calculation that usually requires a full relativistic perturbation theory at second order in metric perturbations. It turns out that the well known result computed in the Newtonian limit continues to apply,

as long as the long-wavelength density perturbation is interpreted as in the synchronous-comving gauge. Our result has the important implication that in full General Relativity, physical modulation of short-scale dynamics must scale with the synchronous-comoving gauge density constrast, rather than the metric perturbation itself. Therefore, our analysis proves, as opposed to previous claims in the literature, that in standard single-clock inflation scenarios, zero scale-dependent galaxy bias is predicted on all scales. This result is crucial to the interpretation of the data from future large-scale structure surveys aiming at detecting primordial non-Gaussianity from new inflationary physics.

For reader's convenience, we collect supporting mathematical details in Appendices.

Chapter 2

Total-Angular-Momentum Wave Formalism

Cosmological perturbations are described as continuous fields with spatial variation and temporal evolution. Some of them are scalar-valued fields, such as the density for the dark matter or the temperature of the CMB. Others are vectorial or tensorial fields that carry multiple components that transform nicely under three-dimensional spatial rotation. The vortical velocity field of the dark matter and the primordial magnetic field are good examples of vector perturbation fields. A famous example for tensor fields is gravitational wave.

We will restrict our discussion to fields living in three spatial dimensions. To avoid confusion with the magnitude of a Fourier wavenumber $k = |\mathbf{k}|$, we use letters $a, b, c, ... = 1, 2, 3$ for tensorial indices. We will use ∇_a for spatial gradient, δ_{ab} for the symmetric

Kronecker delta, and ϵ_{abc} for the anti-symmetric Levi-Civita tensor. If confusion does not arise, we will rely on the type of the argument (real-space $\mathbf{x}$ or Fourier space $\mathbf{k}$) to distinguish between a field in real space and its Fourier transformation. Besides, $(\cdots)$ encloses a pair of symmetrized indices, and $[\cdots]$ a pair of anti-symmetrized indices.

A scalar field $\phi(\mathbf{x})$ has one degree of freedom. A three-dimensional vector field $V^a(\mathbf{x})$ can be decomposed into a transverse component with two degrees of freedom, and a longitudinal component with a single degree of freedom,

$$V^a(\mathbf{x}) \;=\; V^a_T(\mathbf{x}) + V^a_L(\mathbf{x}). \tag{2.1}$$

The transverse component is divergence-free $\nabla_a V^a_T = 0$. The longitudinal component is curl-free $\epsilon_{abc}\nabla^b V^c_L = 0$, and reduces to the gradient of a scalar field. This *transverse-longitudinal decomposition* for a vector field can be generalized to the *scalar-vector-tensor decomposition* (SVT) for a rank-two symmetric tensor field $h^{ab}(\mathbf{x})$ in three spatial dimensions,

$$h^{ab}(\mathbf{x}) \;=\; \frac{1}{3}\delta^{ab}\,h(\mathbf{x}) + h^{ab}_L(\mathbf{x}) + h^{ab}_{TV}(\mathbf{x}) + h^{ab}_{TT}(\mathbf{x}). \tag{2.2}$$

Apart from a scalar trace $h(\mathbf{x})$, the longitudinal scalar part encodes another scalar degree of freedom $h^{ab}_L(\mathbf{x}) \propto (\nabla^a\nabla^b - \delta^{ab}\nabla^2/3)\xi(\mathbf{x})$. The transverse vector part $h^{ab}_{TV}(\mathbf{x})$, obeying $h^{ab}_{TV} \propto \partial^{(a}w^{b)}$ with $\nabla_a w^a = 0$, contains two vectorial degrees of freedom. The transverse, trace-free tensor part $h^{ab}_{TT}(\mathbf{x})$ contains two genuine tensorial degrees of freedom and satisfies $\nabla_a h^{ab}_{TT} = 0$.

The dynamics of cosmological perturbations in general obey nonlinear field equations and are therefore complicated. Fortunately, the perturbation amplitudes are often small in practice, and linearized equations for evolution are excellent approximations. In that case, a perturbation field is conveniently decomposed into normal modes, which obey decoupled equations of motion. The choice for normal modes is conventionally determined by symmetries of the physical problem. Since the bulk property of the Universe is homogenous at fixed cosmic time, according to the *Copernicus Principle*, a popular choice for normal modes are plane waves of definite wave vector $\mathbf{k}$, which are solutions to the Helmholtz equation $(\nabla^2 + \mathbf{k}^2)\Psi(\mathbf{x}) = 0$.

This Chapter is mainly based on the published work of Ref.[42] We first briefly review the plane wave decomposition in Section 2.1. We then start with a spherical wave decomposition for scalar fields in Section 2.2. Generalization to vector fields will be presented in Section 2.3, and to tensor fields in Section 2.4. Concluding remarks will be given in Section 2.5.

2.1 Fourier decompostion

A scalar field can be expanded in terms of plane waves of fixed wave vector $\mathbf{k}$

$$\phi(\mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\phi(\mathbf{k})\Psi^{\mathbf{k}}(\mathbf{x}), \qquad \text{where} \quad \phi(\mathbf{k}) = \int d^3\mathbf{x}\,\phi(\mathbf{x})\left[\Psi^{\mathbf{k}}(\mathbf{x})\right]^*. \tag{2.3}$$

The mode function for scalar plane wave is simply $\Psi^{\mathbf{k}}(\mathbf{x}) = \exp(i\mathbf{k}\cdot\mathbf{x})$, normalized such that

$$\int d^3\mathbf{x}\, \Psi^{\mathbf{k}}(\mathbf{x}) \left[\Psi^{\mathbf{k}'}(\mathbf{x})\right]^* = (2\pi)^3 \delta_D(\mathbf{k}-\mathbf{k}'). \tag{2.4}$$

When generalizing to a vector field, three independent mode functions exist for given $\mathbf{k}$,

$$\begin{aligned}
\Psi_a^{L,\mathbf{k}}(\mathbf{x}) &= (1/k)\nabla_a \Psi^{\mathbf{k}}(\mathbf{x}) = i\hat{k}_a e^{i\mathbf{k}\cdot\mathbf{x}}, \\
\Psi_a^{1,\mathbf{k}}(\mathbf{x}) &= \frac{1}{|\mathbf{k}\times\hat{\mathbf{z}}|}\left(\nabla\times\hat{\mathbf{z}}\Psi^{\mathbf{k}}(\mathbf{x})\right)_a = \frac{i\,[\mathbf{k}\times\hat{\mathbf{z}}]_a}{|\mathbf{k}\times\hat{\mathbf{z}}|} e^{i\mathbf{k}\cdot\mathbf{x}}, \\
\Psi_a^{2,\mathbf{k}}(\mathbf{x}) &= \frac{-i}{k}\epsilon_{ab}{}^c\, \nabla^b \Psi_c^{1,\mathbf{k}}(\mathbf{x}) = \frac{i\,[\mathbf{k}\times(\mathbf{k}\times\hat{\mathbf{z}})]_a}{k|\mathbf{k}\times\hat{\mathbf{z}}|} e^{i\mathbf{k}\cdot\mathbf{x}},
\end{aligned} \tag{2.5}$$

According to the transverse-longitudinal decomposition, two transverse modes have their polarization vectors perpendicular to $\mathbf{k}$, and one longitudinal mode has its polarization vector aligned with $\mathbf{k}$. The vector mode functions obey similar orthonormal relations

$$\int d^3\mathbf{x}\, \Psi_a^{\alpha,\mathbf{k}}(\mathbf{x}) \left[\Psi^{\alpha',\mathbf{k}',a}(\mathbf{x})\right]^* = (2\pi)^3 \delta_{\alpha\alpha'}\delta_D(\mathbf{k}-\mathbf{k}'), \tag{2.6}$$

where $\alpha, \alpha' = L, 1, 2$ run over the three polarization states.

For a rank-2 symmetric tensor field, each Fourier component of the tensor field can be expanded as $\tilde{h}_{ab}(\mathbf{k}) = \sum_s \varepsilon^s_{ab}(\hat{\mathbf{k}}) h_s(\mathbf{k})$ in terms of six polarization states $\varepsilon^s_{ab}(\mathbf{k})$, where $s = \{0, z, x, y, +, \times\}$, for the trace, longitudinal, two vector, and two transverse-traceless polarizations, respectively, with amplitudes denoted by $h_s(\mathbf{k})$.[43] The polarization tensors satisfy $\varepsilon^{s\,ab}\varepsilon^{s'}_{ab} = 2\delta_{ss'}$. The trace polarization tensor is $\varepsilon^0_{ab} \propto \delta_{ab}$, and the longitudinal is

$\varepsilon^z_{ab} \propto (k^a k^b - k^2\delta_{ab}/3)k^{-2}$. The two vector-mode polarization tensors satisfy $\varepsilon^{x,y}_{ab} \propto k_{(a}w^{x,y}_{b)}$ where $w^{x,y}_a$ are two orthogonal ($w^{x\,a}w^y_a = 0$) and transverse ($k^a w^{x,y}_a = 0$) vectors. The two transverse-traceless polarization states have $k^a\varepsilon^{+,\times}_{ab} = 0$. The two vector (spin-1) modes x, y can alternatively be written in terms of a helicity basis by defining two helicity-1 polarization tensors $\varepsilon^{\pm1}_{ab} = (\varepsilon^x_{ab} \pm i\varepsilon^y_{ab})/\sqrt{2}$. Similarly, the two transverse-traceless (spin-2) modes $+, \times$ can alternatively be written in terms of a helicity basis by defining two helicity-2 polarization tensors $\varepsilon^{\pm2}_{ab} = (\varepsilon^+_{ab} \pm i\varepsilon^\times_{ab})/\sqrt{2}$. Analogous to the case of vector field, the tensor plane waves $\Psi^{s,\mathbf{k}}_{ab}(\mathbf{x}) = \varepsilon^s_{ab}(\hat{\mathbf{k}})e^{i\mathbf{k}\cdot\mathbf{x}}$ are normalized to

$$\int d^3\mathbf{x}\, \Psi^{s,\mathbf{k}}_{ab}(\mathbf{x}) \left[\Psi^{s',\mathbf{k}',ab}(\mathbf{x})\right]^* = (2\pi)^3\delta_{ss'}\delta_D(\mathbf{k}-\mathbf{k}'). \tag{2.7}$$

Although commonly adopted in the calculation of cosmological perturbations, the Fourier decomposition has drawbacks. First, translational invariance of plane waves does not necessarily bring about simplification, as cosmological observations are usually carried out on the spherical sky with manifest rotational invariance. In practice, only the past light cone is accessible to the observer, and physics necessarily vary with increased radial distance from the observer, since the further the observer looks, the earlier she traces back in time. Second, for a single Fourier mode, the wavevector $\mathbf{k}$ picks out a preferred direction in the sky, and the plane wave does not transform in a simple way when a spatial rotation about the observer (the spatial coordinate origin) is performed. In particular, a definite Fourier wavelength does not correspond to a fixed angular scale on the sky; conversely, a spherical harmonic observable at a given angular scale receives contribution from an infinite num-

ber of Fourier waves with different wavelengths, resulting in non-trivial convolutions when Fourier power spectra are being converted into spherical-harmonic angular power spectra. The inconvenience becomes the most severe, when vectorial or tensorial observables are involved, as the non-trivial transformation of the polarizaton under spatial rotation adds to the algebraic sophistication.

To address the aforementioned problem, we are motivated to develop the *total-angular-momentum wave formalism* (TAM),[42] a spherical-wave expansion scheme. In this formalism, we preserve manifest rotational symmetry by decomposing perturbation fields into spherical-wave mode functions of definite wave number k [1] (as oppose to a wavevector $\mathbf{k}$), definite *total* angular momentum J and its third component M along a chosen quantization direction, and possible polarization type in the case of vector and tensor fields. The wave number k is inversely proportional to the length scale of variation for the spherical wave. The total angular momentum quantum numbers, J and M, determine the pattern of angular variation across the sky. In the case of vector or tensor spherical waves, total angular momentum will be the sum of "orbital" angular momentum, corresponding to transforming the spatial position $\vec{x}$, and the intrinsic spin, corresponding to transforming the vectorial or tensorial components. In addition, the polarization type describes transversality or handedness.

Apart from the above geometrical consideration, dynamical consideration further justifies the use of spherical-wave expansion in many applications. In linear perturbation theory, translational invariance of physics guarantees that individual Fourier waves are normal

[1] This means that the TAM mode function is solution to the scalar Helmholtz equation $(\nabla^2 + \mathbf{k}^2)\Psi = 0$ or its vector or tensor analogues. Since in cosmological applications we require that mode function be regular at all points, including the spatial origin, we pick up only one solution of the second-order equation.

modes, i.e. they are decoupled from each other. A scheme of spherical-wave expansion still has the same advantage, because normally physical systems in cosmological context are also rotationally invariant. Therefore, TAM waves is an alternative, but desirable choice of normal modes. Beyond linear order in perturbation, study of the general form of cosmological bispectra under the TAM basis has been attempted.[44]

In the following sections, we elaborate on the TAM formalism, distinguishing between the cases of scalar, vector, and tensor fields. We develop an elegant operator formalism to relate TAM wavefunctions of scalar, vector and tensor types.

2.2 TAM decomposition for scalar fields

Since scalar fields carry zero spin, the total angular momentum can be identified with the usual "orbital" angular momentum. Therefore, the TAM expansion for scalar fields are the well-known Fourier-Bessel expansion.

A TAM mode function for scalar fields, having wave number k and angular momentum J and M, has the following factorized form,

$$\Psi^k_{(JM)}(\mathbf{x}) \equiv j_J(kr)Y_{(JM)}(\hat{\mathbf{n}}), \tag{2.8}$$

where $0 < k < +\infty$, $J = 0, 1, 2, \cdots$ and $M = -J, -J+1, \cdots, J-1, J$. Here $j_J(x)$ is a spherical Bessel function of order J, and $Y_{(JM)}(\hat{\mathbf{n}})$ is a spherical harmonic. The spatial position $\mathbf{x}$ is decomposed into a radial distance and a unit vector of direction $\mathbf{x} = r\hat{\mathbf{n}}$. Any

scalar field admits the following expansion,

$$\phi(\mathbf{x}) = \sum_{JM} \int \frac{k^2\, dk}{(2\pi)^3} \phi_{(JM)}(k) 4\pi i^J \Psi^k_{(JM)}(\mathbf{x}), \tag{2.9}$$

with an inverse transformation

$$\phi_{(JM)}(k) = \int d^3\mathbf{x} \left[4\pi i^J \Psi^k_{(JM)}(\mathbf{x})\right]^* \phi(\mathbf{x}) = \int d^2\hat{\mathbf{k}}\, \tilde{\phi}(\mathbf{k}) Y^*_{(JM)}(\hat{\mathbf{k}}). \tag{2.10}$$

The mode functions are orthonormal,

$$(4\pi)^2 \int d^3\mathbf{x} \left[\Psi^k_{(JM)}(\mathbf{x})\right]^* \Psi^{k'}_{(J'M')}(\mathbf{x}) = \delta_{JJ'}\delta_{MM'} \frac{(2\pi)^3}{k^2} \delta_D(k-k'), \tag{2.11}$$

and form a complete basis in the sense that

$$\sum_{JM} \int \frac{k^2\, dk}{(2\pi)^3} \left[4\pi i^J \Psi^k_{(JM)}(\mathbf{x})\right]^* \left[4\pi i^J \Psi^k_{(JM)}(\mathbf{x}')\right] = \delta_D(\mathbf{x}-\mathbf{x}'), \tag{2.12}$$

Fundamental relations Eqs. (2.8)–(2.12) form the basis for computation with a TAM expansion for scalar fields. Expanding a scalar plane wave by scalar TAM waves reproduces the familiar Rayleigh formula

$$e^{i\mathbf{k}\cdot\mathbf{x}} = \sum_{JM} 4\pi i^J Y^*_{(JM)}(\hat{\mathbf{k}}) \Psi^k_{(JM)}(\hat{\mathbf{n}}). \tag{2.13}$$

In cosmology, the initial condition for a perturbation field is often stochastic. The simplest measure of its statistical property is the power spectrum, the Fourier transform of the

2-point correlation function (2PCF). A statistically homogeneous and isotropic perturbation field admits the following form for 2PCF,

$$\left\langle \phi(\mathbf{k})\phi^*(\mathbf{k}')\right\rangle = (2\pi)^3\delta_D(\mathbf{k}-\mathbf{k}')P(k), \tag{2.14}$$

where $P(k)$ is the (isotropic) power spectrum. Here we use $\langle\cdots\rangle$ to refer to ensemble average. When TAM expansion is used, we instead use

$$\left\langle \phi_{(JM)}(k)\phi^*_{(J'M')}(k')\right\rangle = \frac{(2\pi)^3}{k^2}\delta_D(k-k')\delta_{JJ'}\delta_{MM'}P(k). \tag{2.15}$$

It can be seen that for a statistically homogeneous and isotropic random field, the 2PCF under the TAM basis is also diagonal. The relevant power spectrum $P(k)$ coincides with the usual one under the Fourier basis.

2.3 TAM decomposition for vector fields

We now develop TAM basis functions for spin-1 vector fields.

2.3.1 Vector TAM waves

Vector fields carry one spatial index and has spin-1 under rotations. To construct vector TAM waves, the following spherical basis of vectors $e_a^{\bar{m}}$, $\bar{m} = 0, \pm 1$ is a useful notation [2]

$$e_a^0 = e_a^z, \quad e_a^{\pm} = \mp \left(e_a^x \pm i e_a^y\right)/\sqrt{2}. \tag{2.16}$$

We can then straightforwardly construct an *orbital angular momentum basis (OAM)* by adding up orbital angular momentum and spin using the Clebsch-Gordan rule[45]

$$\Psi^{l,k}_{(JM)a}(\mathbf{x}) = j_l(kr) Y^l_{(JM)a}(\hat{\mathbf{n}}) \equiv \sum_{m\bar{m}} \langle 1\bar{m} lm | JM\rangle j_l(kr) Y_{(lm)}(\hat{\mathbf{n}}) e_a^{\bar{m}}. \tag{2.17}$$

Here, Clebsch-Gordan coefficients are given in the notation $\langle l_1 m_1 l_2 m_2 | JM\rangle$. The quantum numbers have range $J = 1, 2, \ldots$, $M = -J, -J+1, \ldots, J-1, J$, and $l = J-1, J, J+1$. These vector basis functions are eigenfunctions of the orbital angular momentum squared operator $\mathbf{L}^2 = L^a L_a$ with eigenvalue $l(l+1)$. They factorize into a radial spherical-Bessel function multiplying an angular vector spherical harmonic $Y^l_{(JM)a}(\hat{\mathbf{n}})$ (of fixed orbital angular momentum l). Following the orthonormal relation for vector sphericla harmonics

$$\int d^2\hat{\mathbf{n}} \left[Y^{l'\,a}_{(J'M')}(\hat{\mathbf{n}})\right]^* Y^l_{(JM)a}(\hat{\mathbf{n}}) = \delta_{ll'}\delta_{JJ'}\delta_{MM'}, \tag{2.18}$$

[2] Barred indices like $\bar{m}$ are reserved for the order-1 spherical basis.

the TAM basis functions satisfy orthonormal relation

$$\int d^3\mathbf{x}\left[4\pi i^J \Psi^{l,k}_{(JM)a}(\mathbf{x})\right]^* 4\pi i^{J'} \Psi^{l',k'\,a}_{(J'M')}(\mathbf{x}) = \delta_{ll'}\delta_{JJ'}\delta_{MM'}\frac{(2\pi)^3}{k^2}\delta_D(k-k'). \tag{2.19}$$

They form a complete functional basis for vector fields,

$$\sum_{JMl}\int \frac{k^2 dk}{(2\pi)^3}\left[4\pi i^l \Psi^{l,k\,a}_{(JM)}(\mathbf{x}')\right]^*\left[4\pi i^l \Psi^{l,k}_{(JM)a}(\mathbf{x})\right] = \delta_D(\mathbf{x}-\mathbf{x}'). \tag{2.20}$$

However, the OAM basis is often inconvenient, because transverse and longitudinal fields are mixed. To improve, *the transverse/longtudinal basis* can be constructed through a linear transformation [3],

$$\begin{aligned}
\Psi^L_{(JM)a}(\mathbf{x}) &= i\left[\sqrt{\frac{J}{2J+1}}\Psi^{J-1}_{(JM)a}(\mathbf{x}) + \sqrt{\frac{J+1}{2J+1}}\Psi^{J+1}_{(JM)a}(\mathbf{x})\right], \\
\Psi^B_{(JM)a}(\mathbf{x}) &= i\Psi^J_{(JM)a}(\mathbf{x}), \\
\Psi^E_{(JM)a}(\mathbf{x}) &= i\left[\sqrt{\frac{J+1}{2J+1}}\Psi^{J-1}_{(JM)a}(\mathbf{x}) - \sqrt{\frac{J}{2J+1}}\Psi^{J+1}_{(JM)a}(\mathbf{x})\right],
\end{aligned} \tag{2.21}$$

Here $\Psi^L_{(JM)a}(\mathbf{x})$ is curl-free, while $\Psi^B_{(JM)a}(\mathbf{x})$, $\Psi^E_{(JM)a}(\mathbf{x})$ are divergence-free, but with opposite parity. Unitarity of the transformation between OAM basis and the transverse/longitudinal basis ensures that the latter are also orthogonal and properly normalized.

The vector TAM basis functions are in fact generated by acting with a set of irreducible vector operators on scalar TAM basis functions. We introduce three irreducible vector

[3] Unless confusion can arise, we will hereafter suppress the label for the wavenumber k, but only explicitly write out the angular quantum numbers. This is because we can restrict to a subset of the Hilbert space with a fixed value of k.

operators,

$$D_a \equiv (i/k)\nabla_a, \qquad K_a \equiv -iL_a, \qquad M_a \equiv \epsilon_{abc} D^b K^c. \tag{2.22}$$

Since they all commute with ∇^2, these expressions have been made specific to a sub-space of definite eigenvalue $\nabla^2 = k^2$. Since they form an irreducible vector representation, they conserve total angular momentum when generating a vector field from a scalar function. The three operators satisfy orthogonal relations

$$D^a K_a = K^a D_a = 0, \qquad D^a M_a = 0, \qquad M^a D_a = 2, \qquad K^a M_a = M^a K_a = 0, \tag{2.23}$$

and their hermitian conjugates are given by

$$(D_a)^\dagger = D_a, \qquad (K_a)^\dagger = -K_a, \qquad (M_a)^\dagger = -M_a + 2D_a. \tag{2.24}$$

Those three operators obey the following commutation relations

$$\begin{aligned} &[D_a, D_b] = 0, \qquad [K_a, D_b] = \epsilon_{abc} D^c, \qquad [M_a, D_b] = g_{ab} - D_a D_b, \\ &[K_a, K_b] = \epsilon_{abc} K^c, \qquad [K_a, M_b] = \epsilon_{abc} M^c, \qquad [M_a, M_b] = -\epsilon_{abc} K^c. \end{aligned} \tag{2.25}$$

In particular, a subset of those operators K_a and M_a form a closed Lorentz algebra.

When acting on the scalar TAM wave $\Psi_{(JM)}(\mathbf{x})$, the operators D_a, K_a and M_a have a simple limiting behavior in the far-field limit $kr \to \infty$, for which the operator D_a asymptotes

to the radial direction, and operators K_a and M_a both become orthogonal to the radial direction, with K_a obtained by a $+90^\circ$ rotation of M_a about the outward radial direction. From this construction, the longitudinal vector TAM mode is generated by D_a, and the two transverse vector modes, E mode and B mode, are generated by K_a and M_a, respectively,

$$\begin{aligned}
\Psi^L_{(JM)a}(\mathbf{x}) &= D_a \Psi_{(JM)}(\mathbf{x}), \\
\Psi^B_{(JM)a}(\mathbf{x}) &= \frac{K_a}{\sqrt{J(J+1)}} \Psi_{(JM)}(\mathbf{x}), \\
\Psi^E_{(JM)a}(\mathbf{x}) &= \frac{M_a}{\sqrt{J(J+1)}} \Psi_{(JM)}(\mathbf{x}).
\end{aligned} \tag{2.26}$$

It is worth to point out, however, that at finite radius r, $\Psi^L_{(JM)a}(\mathbf{x})$ and $\Psi^E_{(JM)a}(\mathbf{x})$ have both radial and transverse components, while $\Psi^B_{(JM)a}(\mathbf{x})$ is purely perpendicular to the radial direction. By "transverse" or "longitudinal", we are referring to whether a vector field is a gradient field or a curl field, rather than referring to the point-by-point orientation of the vector field with respect to the radial direction, although these two interpretations coincide in the limit of infinite radius.

The helicity basis is especially convenient for discussing parity violating physics. They are eigenfunctions of the helicity operator, defined as $H = \mathbf{S} \cdot \hat{\mathbf{p}}$, where we introduce the spin operator $(S_b)_{ac} = i\epsilon_{abc}$, and the normalized momentum operator $\hat{p}_a = -i\nabla_a / k$. Three integer eigenvalues for the helicity operator $\lambda = 0, \pm 1$ are possible. It turns out the longitudinal TAM wave has helicity zero $\lambda = 0$, and the transverse TAM waves are linear combinations

of unity helicity $\lambda = \pm 1$ eigenmodes,

$$\Psi^{\pm 1}_{(JM)a}(\mathbf{x}) = \frac{1}{\sqrt{2}}\left[\Psi^{E}_{(JM)a}(\mathbf{x}) \pm i\Psi^{B}_{(JM)a}(\mathbf{x})\right], \qquad \Psi^{0}_{(JM)a}(\mathbf{x}) = \Psi^{L}_{(JM)a}(\mathbf{x}). \tag{2.27}$$

It is clear that the helicity basis also separates the longitudinal component and the transverse component.

In summary, we have constructed three distinct basis for vector TAM waves — OAM basis $\Psi^{k,l}_{(JM)a}(\mathbf{x})$ with $l = J, J\pm 1$ (Eq. (2.17)), the transverse/longitudinal basis (Eq. (2.21)), and the helicity basis (Eq. (2.27)). The first basis are straightforwardly constructed using the Clebsch-Gordan rule, and the second and third basis distinguish between longitudinal and transverse vector fields.

2.3.2 Projection onto vector spherical harmonics

The aforementioned construction for vector TAM basis functions linearly express all vector fields as a function of three dimensional position $\vec{x}$. In applications involving spherical harmonic observables on the two-dimensional sky — vector spherical harmonics $Y^{\alpha}_{(JM)a}(\hat{\mathbf{n}})$, vector fields that live on a two dimensional unit sphere but can point either in the radial direction or tangent to the sphere — have been previously introduced.[46] These are independent linear combinations of vector spherical harmonics of fixed orbital angular momentum l, as defined in Eq. (2.17). These provide a complete set of basis functions for three-dimensional vector fields that are only dependent on the direction in the sky $\hat{\mathbf{n}}$.

Similar to the construction of vector TAM waves, vector spherical harmonics can be

generated by differential operators acting on the usual scalar spherical harmonics. Let us introduce three operators

$$N_a = -\hat{n}_a, \qquad K_a = -iL_a, \qquad M_{\perp a} = \epsilon_{abc} N^b K^c. \tag{2.28}$$

These are dimensionless operators that only act on the angular direction $\hat{\mathbf{n}}$, but not on the radius r [4]. These are analogous to the three operators in Eq. (2.22), in the sense that algebraic relations Eqs. (2.23)–(2.25) can be carried through, once the three operators D_a, K_a and M_a are replaced with three new operators N_a, K_a and $M_{\perp a}$, respectively. The two transverse operators K_a and $M_{\perp a}$ generate the two vector spherical harmonics $Y^E_{(JM)a}(\hat{\mathbf{n}})$ and $Y^B_{(JM)a}(\hat{\mathbf{n}})$ perpendicular to the line of sight, while the longitudinal N_a generates a third vector spherical harmonic $Y^L_{(JM)a}(\hat{\mathbf{n}})$ along the line of sight,

$$\begin{aligned}
Y^B_{(JM)a}(\hat{\mathbf{n}}) &= \frac{1}{\sqrt{J\,(J+1)}} K_a Y_{(JM)}(\hat{\mathbf{n}}) = i Y^J_{(JM)a}(\hat{\mathbf{n}}), \\
Y^E_{(JM)a}(\hat{\mathbf{n}}) &= \frac{1}{\sqrt{J\,(J+1)}} M_{\perp a} Y_{(JM)}(\hat{\mathbf{n}}) = -\sqrt{\frac{J+1}{2J+1}} Y^{J-1}_{(JM)a}(\hat{\mathbf{n}}) - \sqrt{\frac{J}{2J+1}} Y^{J+1}_{(JM)a}(\hat{\mathbf{n}}), \\
Y^L_{(JM)a}(\hat{\mathbf{n}}) &= N_a Y_{(JM)}(\hat{\mathbf{n}}) = -\sqrt{\frac{J}{2J+1}} Y^{J-1}_{(JM)a}(\hat{\mathbf{n}}) + \sqrt{\frac{J+1}{2J+1}} Y^{J+1}_{(JM)a}(\hat{\mathbf{n}}),
\end{aligned} \tag{2.29}$$

satisfying the orthonormal relation

$$\int d^2\hat{\mathbf{n}} \left[Y^{\alpha,a}_{(JM)}(\hat{\mathbf{n}}) \right]^* Y^{\beta}_{(J'M')a}(\hat{\mathbf{n}}) = \delta_{JJ'} \delta_{MM'} \delta_{\alpha\beta}. \tag{2.30}$$

where $\alpha, \beta = \{E, B, L\}$. At given direction $\hat{\mathbf{n}}$, the three vector spherical harmonics are

[4] Note that K_a in Eq. (2.28) is the same differential operator as K_a in Eq. (2.22).

perpendicular to each other.

Although the $E/B/L$ vector TAM waves are not factorizable into a radial profile multiplied by an angular function (e.g. vector spherical harmonics), they can be expanded in terms of vector spherical harmonics of Eq. (2.29),

$$\begin{aligned}
\Psi^{k,B}_{(JM)a}(\mathbf{x}) &= j_J(kr)Y^B_{(JM)a}(\hat{\mathbf{n}}), \\
\Psi^{k,E}_{(JM)a}(\mathbf{x}) &= -i\left[j'_J(kr)+\frac{j_J(kr)}{kr}\right]Y^E_{(JM)a}(\hat{\mathbf{n}}) - i\sqrt{J(J+1)}\frac{j_J(kr)}{kr}Y^L_{(JM)a}(\hat{\mathbf{n}}), \\
\Psi^{k,L}_{(JM)a}(\mathbf{x}) &= -i\sqrt{J(J+1)}\frac{j_J(kr)}{kr}Y^E_{(JM)a}(\hat{\mathbf{n}}) - ij'_J(kr)Y^L_{(JM)a}(\hat{\mathbf{n}}),
\end{aligned} \tag{2.31}$$

where the expansion coefficients, dependent on the radial variable, are related to the usual spherical Bessel function and its derivatives. These results are useful when physical fields are decomposed into a line-of-sight component and a perpendicular component.

2.3.3 TAM expansion of vector plane wave

For conversion between Fourier expansion and the TAM expansion, we seek spherical-wave expansion of a vector plane wave with wave vector $\mathbf{k}$ and polarization vector $\hat{\varepsilon}_a(\mathbf{k})$, an analogue of the Rayleigh formula in the scalar case.

For the OAM basis, the expansion reads

$$\hat{\varepsilon}_a(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}} = \sum_{lJM} 4\pi i^l A^l_{(JM)}(\hat{\mathbf{k}})\Psi^{l,k}_{(JM)a}(\mathbf{x}), \tag{2.32}$$

where the expansion coeffecients are given by

$$A^l_{(JM)}(\hat{\mathbf{k}}) = \hat{\varepsilon}^a(\mathbf{k}) Y^{l\,*}_{(JM)a}(\hat{\mathbf{k}}). \tag{2.33}$$

If one chooses to use the $E/B/L$ basis, or the helicity basis, the expansion admits a form of similar structure

$$\hat{\varepsilon}_a(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} = \sum_{\alpha=L,E,B} \sum_{JM} 4\pi i^J A^\alpha_{(JM)}(\hat{\mathbf{k}}) \Psi^{k,\alpha}_{(JM)a}(\mathbf{x}), \tag{2.34}$$

$$\hat{\varepsilon}_a(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} = \sum_{\lambda=-1,0,1} \sum_{JM} 4\pi i^J A^\lambda_{(JM)}(\hat{\mathbf{k}}) \Psi^{k,\lambda}_{(JM)a}(\mathbf{x}), \tag{2.35}$$

with expansion coefficients

$$A^\alpha_{(JM)}(\hat{\mathbf{k}}) = \hat{\varepsilon}^a(\mathbf{k}) Y^{\alpha\,*}_{(JM)a}(\hat{\mathbf{k}}), \tag{2.36}$$

$$A^\lambda_{(JM)}(\hat{\mathbf{k}}) = \hat{\varepsilon}^a(\mathbf{k}) Y^{\lambda\,*}_{(JM)a}(\hat{\mathbf{k}}). \tag{2.37}$$

For the helicity basis, we have introduced the helical vector spherical harmonics $Y^{\lambda=\pm1}_{(JM)a} = 2^{-1/2}\left[Y^E_{(JM)a} \pm iY^B_{(JM)a}\right]$, which are closely related to the familiar spin-1 spherical harmonics[47–50] through

$$\hat{\varepsilon}^a_{\lambda'}(\mathbf{k}) Y^\lambda_{(JM)a}(\hat{\mathbf{k}}) = {}_{-\lambda}Y_{(JM)}(\hat{\mathbf{k}}) \delta_{\lambda\lambda'}, \qquad \lambda, \lambda' = 0, \pm 1. \tag{2.38}$$

Our convention for the polarization vectors $\hat{\varepsilon}^a_\lambda(\mathbf{k})$ for $\lambda = 0, \pm 1$, defined with respect to

given wave vector $\mathbf{k}$, is given by

$$\hat{\varepsilon}^a_0 = \hat{n}^a, \qquad \hat{\varepsilon}^a_{\pm 1} = \mp\frac{1}{\sqrt{2}}(\hat{\theta}^a \mp i\hat{\phi}^a). \tag{2.39}$$

2.3.4 Expansion of vector fields and power spectra

Any vector field $V_a(\mathbf{x})$ admits an expansion in the $E/B/L$ vector TAM waves

$$V_a(\mathbf{x}) = \sum_{JM}\sum_{\alpha=E,B,L}\int \frac{k^2 dk}{(2\pi)^3} V^l_{(JM)}(k) 4\pi i^l \Psi^{\alpha,k}_{(JM)a}(\mathbf{x}), \tag{2.40}$$

with the inversion

$$V^\alpha_{(JM)}(k) = \int d^3\mathbf{x}\, V^a(\mathbf{x}) \left[4\pi i^J \Psi^{\alpha,k}_{(JM)a}(\mathbf{x})\right]^*. \tag{2.41}$$

Similar relations hold for the OAM basis or the helicity basis. When a stochastic vector field has statistically homogeneous *and* isotropic 2-point correlation functions, the E/B TAM components must share the same power spectrum, while it may differ from that of the L TAM component, i.e.

$$\left\langle \left[V^\alpha_{(JM)}(k)\right]^* V^\beta_{(J'M')}(k') \right\rangle = P_T(k)\delta_{JJ'}\delta_{MM'}\delta_{\alpha\beta}\frac{(2\pi)^3}{k^2}\delta_D(k-k'), \tag{2.42}$$

for $\{\alpha,\beta\} = \{E,B\}$, and

$$\left\langle \left[V^L_{(JM)}(k)\right]^* V^L_{(J'M')}(k') \right\rangle = P_L(k)\delta_{JJ'}\delta_{MM'}\frac{(2\pi)^3}{k^2}\delta_D(k-k'). \tag{2.43}$$

The longitudinal mode does not necessarily share the same statistics as the E/B transverse modes because it is essentially the gradient of a scalar degree of freedom. Note that an unequal power between the E/B transverse modes necessarily signals a violation of statistical homogeneity and implies a preferred center in the Universe,[51] while a nonzero cross-correlation between the E/B transverse modes indicates a violation of parity.

2.4 TAM decomposition for tensor fields

We now develop TAM basis functions for spin-2 tensor fields, which find an abundance of applications in the study of cosmological perturbations. The presentation will largely parallel our discussion on vector fields in Section 2.3. Similar studies have existed in the literature of gravitational radiation,[52] although compact expressions for three-dimensional TAM mode functions for tensor fields had not been explicitly written down prior to the work of Ref.[42]

2.4.1 Tensor TAM waves

A rank-two symmetric tensor-valued field $h_{ab}(\mathbf{x})$ has two spatial indices and carries an intrinsic spin of two units. It has a total of six degrees of freedom. Since the trace part transforms trivially under spatial rotations as a spin-zero scalar field, for the remainder of our discussion, we restrict to trace-free tensor field $\delta^{ab}h_{ab}(\mathbf{x}) = 0$, which furnishes a single irreducible representation under rotation.

To construct TAM waves for tensor fields, we first construct five spin-2 spherical basis

tensors $t^{\tilde{m}}_{ab}$, $\tilde{m} = 0, \pm 1, \pm 2$,

$$t^{\tilde{m}}_{ab} \equiv \sum_{\bar{m}_1 \bar{m}_2} \langle 1\bar{m}_1 1\bar{m}_2 | 2\tilde{m}\rangle e^{\bar{m}_1}_a e^{\bar{m}_2}_b, \tag{2.44}$$

by taking a direct product of two copies of spin-1 spherical basis vectors of Eq. (2.16), through the Clebsch-Gordan rule of addition. These basis tensors are by construction symmetric and trace-free. Entirely analogous to the case of vector fields, an OAM basis of tensor wavefunction can be straightforwardly constructed by "adding up" the orbital angular momentum with the intrinsic spin

$$\Psi^{l,k}_{(JM)ab}(\mathbf{x}) \equiv j_l\,(kr)\, Y^l_{(JM)ab}(\hat{\mathbf{n}}) = \sum_{\tilde{m}m} \langle 2\tilde{m}lm|JM\rangle j_l\,(kr)\, Y_{(lm)}(\hat{\mathbf{n}}) t^{\tilde{m}}_{ab}. \tag{2.45}$$

Note that for tensor basis functions, the total angular momentum quantum number J is equal or larger than two. For each value of J, there are five possible integer values of orbital angular momentum $l = J-2, J-1, J, J+1, J+2$. These tensor basis functions are eigenfunctions of the orbital angular momentum squared operator $\mathbf{L}^2$ with eigenvalue $l(l+1)$. They factorize into a radial spherical-Bessel profile, and a tensor spherical harmonic $Y^l_{(JM)ab}(\hat{\mathbf{n}})$ satisfying

$$\int d^2\hat{\mathbf{n}}\, Y^{l\ ab}_{(JM)}(\hat{\mathbf{n}}) Y^{l'\,*}_{(J'M')ab}(\hat{\mathbf{n}}) = \delta_{ll'}\delta_{JJ'}\delta_{MM'}. \tag{2.46}$$

The OAM basis functions are then normalized such that

$$\int d^3x \left[4\pi i^J \Psi^{l,k}_{(JM)ab}(\mathbf{x})\right]^* 4\pi i^{J'} \Psi^{l',k'\,ab}_{(J'M')}(\mathbf{x}) = \delta_{ll'}\delta_{JJ'}\delta'_{MM}\frac{(2\pi)^3}{k^2}\delta_D(k-k'), \tag{2.47}$$

again analogous to the case of vector fields. These also form a complete functional basis for any symmetric, trace-free tensor field

$$\sum_{JMl}\int \frac{k^2dk}{(2\pi)^3}\left[4\pi i^l\Psi^{l,k\ ab}_{(JM)}(\mathbf{x}')\right]^*\left[4\pi i^l\Psi^{l,k}_{(JM)ab}(\mathbf{x})\right]=\delta_D(\mathbf{x}-\mathbf{x}'). \tag{2.48}$$

However, the OAM basis for tensor fields do not automatically fullfil the SVT decomposition of Eq. (2.2). As an example, a gravitational-wave field possesses only the two divergence-free degrees of freedom, instead of all five degrees of freedom in general. For this reason, we seek a unitary transformation into the *longitudinal/vector/transverse-traceless (LVT) basis*. We first obtain a longitudinal scalar mode,

$$\begin{aligned}\Psi^L_{(JM)ab}(\mathbf{x})=&\left(\frac{3\,(J-1)\,J}{2\,(2J-1)\,(2J+1)}\right)^{1/2}\Psi^{J-2}_{(JM)ab}(\mathbf{x})+\left(\frac{J\,(J+1)}{(2J-1)\,(2J+3)}\right)^{1/2}\Psi^{J}_{(JM)ab}(\mathbf{x})\\&+\left(\frac{3\,(J+1)\,(J+2)}{2\,(2J+1)\,(2J+3)}\right)^{1/2}\Psi^{J+2}_{(JM)ab}(\mathbf{x}),\end{aligned} \tag{2.49}$$

which may be expressed as the trace-free Hesian of a scalar TAM wave $\Psi^L_{(JM)ab}(\mathbf{x})=\frac{1}{k^2}\sqrt{\frac{3}{2}}\left(\nabla_a\nabla_b-\frac{1}{3}\delta_{ab}\nabla^2\right)\Psi_{(JM)}(\mathbf{x})$. We then find two vectorial modes (VE and VB modes)

$$\begin{aligned}\Psi^{VB}_{(JM)ab}(\mathbf{x})=&\left(\frac{J-1}{2J+1}\right)^{1/2}\Psi^{J-1}_{(JM)ab}(\mathbf{x})+\left(\frac{J+2}{2J+1}\right)^{1/2}\Psi^{J+1}_{(JM)ab}(\mathbf{x}),\\\Psi^{VE}_{(JM)ab}(\mathbf{x})=&\left(\frac{2\,(J-1)\,(J+1)}{(2J-1)\,(2J+1)}\right)^{1/2}\Psi^{J-2}_{(JM)ab}(\mathbf{x})+\left(\frac{3}{(2J-1)\,(2J+3)}\right)^{1/2}\Psi^{J}_{(JM)ab}(\mathbf{x})\\&-\left(\frac{2J\,(J+2)}{(2J+1)\,(2J+3)}\right)^{1/2}\Psi^{J+2}_{(JM)ab}(\mathbf{x}).\end{aligned} \tag{2.50}$$

These have divergences equal to the correspondent vector TAM waves $\frac{1}{k}\nabla^a\Psi^{VB}_{(JM)ab}(\mathbf{x})=$

$\frac{i}{\sqrt{2}}\Psi^{B}_{(JM)b}(\mathbf{x})$ and $\frac{1}{k}\nabla^{a}\Psi^{VE}_{(JM)ab}(\mathbf{x}) = \frac{i}{\sqrt{2}}\Psi^{E}_{(JM)b}(\mathbf{x})$, and are in fact gradients of vector TAM waves. Finally, we obtain the two transverse-traceless modes, the "genuine" tensor modes,

$$\begin{aligned}\Psi^{TE}_{(JM)ab}(\mathbf{x}) \equiv & \left(\frac{(J+1)(J+2)}{2(2J-1)(2J+1)}\right)^{1/2}\Psi^{J-2}_{(JM)ab}(\mathbf{x}) - \left(\frac{3(J-1)(J+2)}{(2J-1)(2J+3)}\right)^{1/2}\Psi^{J}_{(JM)ab}(\mathbf{x}) \\ & + \left(\frac{J(J-1)}{2(2J+1)(2J+3)}\right)^{1/2}\Psi^{J+2}_{(JM)ab}(\mathbf{x}), \\ \Psi^{TB}_{(JM)ab}(\mathbf{x}) \equiv & \left(\frac{J+2}{2J+1}\right)^{1/2}\Psi^{J-1}_{(JM)ab}(\mathbf{x}) - \left(\frac{J-1}{2J+1}\right)^{1/2}\Psi^{J+1}_{(JM)ab}(\mathbf{x}),\end{aligned} \tag{2.51}$$

which have vanishing divergence. Eq. (2.49), Eq. (2.50) and Eq. (2.51) together constitute a complete list of five independent degrees of freedom for given total angular momentum J, M and given wavenumber k (which we often suppress for a compact notation), which are compatible with the SVT decomposition. Since the linear transformation coefficients in Eq. (2.49), Eq. (2.50) and Eq. (2.51) are unitary, the LVT basis are orthonormal according to

$$(4\pi)^2\int d^3x\,\left[\Psi^{\alpha,k}_{(JM)ab}(\mathbf{x})\right]^{*}\Psi^{\beta,k\ ab}_{(J'M')}(\mathbf{x}) = \delta_{\alpha\beta}\delta_{JJ'}\delta_{MM'}\frac{(2\pi)^3}{k^2}\delta_D(k-k'), \tag{2.52}$$

for $\{\alpha,\beta\} = \{L, VB, VE, TB, TE\}$.

The LVT TAM basis for tensor fields can be generated from a set of tensor operators acting on the scalar TAM wave. We have proposed three vector operators D_a, K_a and M_a

in Eq. (2.22). Using those as basic building blocks, five tensor operators can be constructed

$$
\begin{aligned}
&T^{L}_{ab} = -D_a D_b + \frac{1}{3}\delta_{ab}, \qquad && T^{VB}_{ab} = D_{(a}K_{b)}, \qquad T^{VE}_{ab} = D_{(a}M_{b)},\\
&T^{TB}_{ab} = K_{(a}M_{b)} + M_{(a}K_{b)} + 2D_{(a}K_{b)}, \qquad && T^{TE}_{ab} = M_{(a}M_{b)} - K_{(a}K_{b)} + 2D_{(a}M_{b)}.
\end{aligned}
\tag{2.53}
$$

These form an irrreducible representation of spin-2 since they are by construction symmetric and trace-free. These generate, when acting on a scalar TAM wave, orthogonal tensor-valued functions because

$$
\left(T^{\alpha'}_{ab}\right)^{\dagger} T^{\alpha,ab} = 0, \quad \text{if } \alpha \neq \alpha', \qquad \text{for} \qquad \alpha, \alpha' = L, VB, VE, TB, TE. \tag{2.54}
$$

For $\alpha = \alpha'$, the numerical value on the right hand side (when acting on a scalar TAM eigenfunction) sets the normalization. These operators then exactly generate the LVT basis of Eq. (2.49), Eq. (2.50) and Eq. (2.51),

$$
\begin{aligned}
&\Psi^{L}_{(JM)ab}(\mathbf{x}) = \sqrt{\frac{3}{2}} T^{L}_{ab}\Psi_{(JM)}(\mathbf{x}), \qquad (2.55)\\
&\Psi^{VB}_{(JM)ab}(\mathbf{x}) = -\sqrt{\frac{2}{J\,(J+1)}} T^{VB}_{ab}\Psi_{(JM)}(\mathbf{x}), \quad \Psi^{VE}_{(JM)ab}(\mathbf{x}) = -\sqrt{\frac{2}{J\,(J+1)}} T^{VE}_{ab}\Psi_{(JM)}(\mathbf{x}),\\
&\Psi^{TB}_{(JM)ab}(\mathbf{x}) = -\sqrt{\frac{(J-2)!}{2(J+2)!}} T^{TB}_{ab}\Psi_{(JM)}(\mathbf{x}), \quad \Psi^{TE}_{(JM)ab}(\mathbf{x}) = -\sqrt{\frac{(J-2)!}{2(J+2)!}} T^{TE}_{ab}\Psi_{(JM)}(\mathbf{x}).
\end{aligned}
$$

It is obvious from the operator construction that although the LVT basis are orthogonal tensor-valued functions, their tensorial values at any finite point are not strictly orthogonal, unless the far-field limit $kr \to \infty$ is taken. Let $\hat{n}^a$ be the radial unit vector. Then in the far-

field limit, the L mode has a "longitudinal-scalar" type polarization tensor $\propto (\hat{n}_a\hat{n}_b - \delta_{ab}/3)$; the vectorial VE/VB modes have polarizations $\propto n_{(a}w^{E,B}_{b)}$, where $w^{E,B}_a$ are two orthogonal vectors $w^E_a w^{B,a} = 0$ perpendicular to the radial direction $\hat{n}^a w^{E,B}_a = 0$; finally, the transverse tensorial TE/TB modes have polarizations completely in the transverse plane. We therefore see that in the far-field limit, the LVT basis locally resembles plane waves propagating along the radial direction (or more precisely standing waves), exactly corresponding to the five plane-wave polarization states categorized in Section 2.1, according to the orientation with respect to the plane wave vector.

$$\Psi^{\pm 2}_{(JM)ab}(\mathbf{x}) = \frac{1}{\sqrt{2}}\left(\Psi^{TE}_{(JM)ab}(\mathbf{x}) \pm i\Psi^{TB}_{(JM)ab}(\mathbf{x})\right),$$
$$\Psi^{\pm 1}_{(JM)ab}(\mathbf{x}) = \frac{1}{\sqrt{2}}\left(\Psi^{VE}_{(JM)ab}(\mathbf{x}) \pm i\Psi^{VB}_{(JM)ab}(\mathbf{x})\right),$$
$$\Psi^{0}_{(JM)ab}(\mathbf{x}) = \Psi^{L}_{(JM)ab}(\mathbf{x}). \tag{2.56}$$

In summary, three basis for tensor TAM waves — OAM basis $\Psi^{l,k}_{(JM)ab}(\mathbf{x})$ with $l = J, J\pm 1, J\pm 2$ (Eq. (2.45)), the longitudinal/vector/transverse-traceless basis (Eq. (2.55)), and the helicity basis (Eq. (2.56)). The first basis are straightforwardly constructed using the Clebsch-Gordan rule, and the second and third basis distinguish between longitudinal and transverse vector fields.

2.4.2 Projection onto tensor spherical harmonics

Three dimensional tensor TAM waves can be expanded in terms of tensor spherical harmonics $Y^{\alpha}_{(JM)ab}(\hat{\mathbf{n}})$, tensor-valued functions living on the two-dimensional sky.[46,53–55] We parallel the operator approach for the case of vector field to reproduce these tensor spherical harmonics. In Eq. (2.28) we have introduced operators $\{N_a, K_a, M_{\perp a}\}$, which act only on angular variables and satisfy the same algebra as $\{D_a, K_a, M_a\}$. We then use those as building blocks and construct five tensor operators which again only act on the angular part of a wavefunction

$$W^{L}_{ab} = -N_a N_b + \frac{1}{3} g_{ab}, \qquad W^{VB}_{ab} = N_{(a} K_{b)}, \qquad W^{VE}_{ab} = N_{(a} M_{\perp b)},$$
$$W^{TB}_{ab} = K_{(a} M_{\perp b)} + M_{\perp (a} K_{b)} + 2 N_{(a} K_{b)}, \qquad W^{TE}_{ab} = M_{\perp (a} M_{\perp b)} - K_{(a} K_{b)} + 2 N_{(a} M_{\perp b)}, \tag{2.57}$$

These symmetric and traceless operators form an irreducible representation of spin-2. They satisfy the same algebraic property as the operators defined in Eq. (2.53) do. Because they conserve total angular momentum, they generate tensor spherical harmonics of fixed total angular momentum J, M when acting on scalar spherical harmonics of the same angular

momentum J, M, which read

$$
\begin{aligned}
Y^{L}_{(JM)ab}(\hat{\mathbf{n}}) &= \sqrt{\frac{3}{2}} W^{L}_{ab} Y_{(JM)}(\hat{\mathbf{n}}), \qquad (2.58)\\
Y^{VE}_{(JM)ab}(\hat{\mathbf{n}}) &= -\sqrt{\frac{2}{J(J+1)}} W^{VE}_{ab} Y_{(JM)}(\hat{\mathbf{n}}), \qquad Y^{VB}_{(JM)ab}(\hat{\mathbf{n}}) = -\sqrt{\frac{2}{J(J+1)}} W^{VB}_{ab} Y_{(JM)}(\hat{\mathbf{n}}),\\
Y^{TE}_{(JM)ab}(\hat{\mathbf{n}}) &= -\sqrt{\frac{(J-2)!}{2(J+2)!}} W^{TE}_{ab} Y_{(JM)}(\hat{\mathbf{n}}), \qquad Y^{TB}_{(JM)ab}(\hat{\mathbf{n}}) = -\sqrt{\frac{(J-2)!}{2(J+2)!}} W^{TB}_{ab} Y_{(JM)}(\hat{\mathbf{n}}),
\end{aligned}
$$

These are linear combinations of the OAM tensor spherical harmonics $Y^{l}_{(JM)ab}(\hat{\mathbf{n}})$ for $l = J, J \pm 1, J \pm 2$ as defined in Eq. (2.45)

$$
\begin{aligned}
Y^{L}_{(JM)ab}(\hat{\mathbf{n}}) &= -\left(\frac{3J\,(J-1)}{2\,(2J-1)\,(2J+1)}\right)^{\frac{1}{2}} Y^{J-2}_{(JM)ab}(\hat{\mathbf{n}}) + \left(\frac{J\,(J+1)}{(2J-1)\,(2J+3)}\right)^{\frac{1}{2}} Y^{J}_{(JM)}(\hat{\mathbf{n}})\\
&\quad -\left(\frac{3\,(J+1)\,(J+2)}{2\,(2J+1)\,(2J+3)}\right)^{\frac{1}{2}} Y^{J+2}_{(JM)}(\hat{\mathbf{n}}),\\
Y^{VE}_{(JM)ab}(\hat{\mathbf{n}}) &= -\left(\frac{2\,(J-1)\,(J+1)}{(2J-1)\,(2J+1)}\right)^{\frac{1}{2}} Y^{J-2}_{(JM)ab}(\hat{\mathbf{n}})\\
&\quad +\left(\frac{3}{(2J-1)\,(2J+3)}\right)^{\frac{1}{2}} Y^{J}_{(JM)ab}(\hat{\mathbf{n}}) + \left(\frac{2J\,(J+2)}{(2J+1)\,(2J+3)}\right)^{\frac{1}{2}} Y^{J+2}_{(JM)ab}(\hat{\mathbf{n}}),\\
Y^{VB}_{(JM)ab}(\hat{\mathbf{n}}) &= i\left(\frac{J-1}{2J+1}\right)^{\frac{1}{2}} Y^{J-1}_{(JM)ab}(\hat{\mathbf{n}}) - i\left(\frac{J+2}{2J+1}\right)^{\frac{1}{2}} Y^{J+1}_{(JM)ab}(\hat{\mathbf{n}}),\\
Y^{TE}_{(JM)ab}(\hat{\mathbf{n}}) &= -\left(\frac{(J+1)\,(J+2)}{2\,(2J-1)\,(2J+1)}\right)^{\frac{1}{2}} Y^{J-2}_{(JM)ab}(\hat{\mathbf{n}}) - \left(\frac{3\,(J-1)\,(J+2)}{(2J-1)\,(2J+3)}\right)^{\frac{1}{2}} Y^{J}_{(JM)ab}(\hat{\mathbf{n}})\\
&\quad -\left(\frac{J\,(J-1)}{2\,(2J+1)\,(2J+3)}\right)^{\frac{1}{2}} Y^{J+2}_{(JM)ab}(\hat{\mathbf{n}}),\\
Y^{TB}_{(JM)ab}(\hat{\mathbf{n}}) &= i\left(\frac{J+2}{2J+1}\right)^{\frac{1}{2}} Y^{J-1}_{(JM)ab}(\hat{\mathbf{n}}) + i\left(\frac{J-1}{2J+1}\right)^{\frac{1}{2}} Y^{J+1}_{(JM)ab}(\hat{\mathbf{n}}). \qquad (2.59)
\end{aligned}
$$

This basis of tensor spherical harmonics are normalized according to

$$
\int d^2\hat{\mathbf{n}} \left[Y^{\alpha,ab}_{(JM)}(\hat{\mathbf{n}})\right]^{*} Y^{\beta}_{(J'M')ab}(\hat{\mathbf{n}}) = \delta_{JJ'}\delta_{MM'}\delta_{\alpha\beta}, \qquad (2.60)
$$

where $\alpha, \beta = \{L, VE, VB, TE, TB\}$. At given direction $\hat{\mathbf{n}}$, the five tensor spherical harmonics can be distinguished by a decomposition into line-of-sight components and perpendicular components. The longitudinal L type has a polarization tensor $\propto (\hat{n}_a \hat{n}_b - \delta_{ab}/3)$. The two vectorial VE/VB types are only partially perpendicular to the line of sight, i.e. $\hat{n}^a \hat{n}^b Y^{TE/TB}_{(JM)ab}(\hat{\mathbf{n}}) = 0$ but $\hat{n}^a Y^{TE/TB}_{(JM)ab}(\hat{\mathbf{n}}) \neq 0$. Finally, the two tensorial TE/TB modes are entirely perpendicular to the line of sight $\hat{n}^a Y^{TE/TB}_{(JM)ab}(\hat{\mathbf{n}}) = 0$; they form a useful basis to expand observables that live on the two-dimensional sky.

We are now in a position to find decomposition of tensor TAM waves $\Psi^{k,\alpha}_{(JM)ab}(\mathbf{x})$ in terms of tensor spherical harmonics of Eq. (2.59). The results read

$$
\begin{aligned}
\Psi^{k,L}_{(JM)ab}(\mathbf{x}) =& -\frac{1}{2}\left(j_J(kr) + 3g_J(kr)\right) Y^{L}_{(JM)ab}(\hat{\mathbf{n}}) - \sqrt{3J(J+1)} f_J(kr)\, Y^{VE}_{(JM)ab}(\hat{\mathbf{n}}) \\
& - \frac{\sqrt{3J(J+1)(J-1)(J+2)}}{2} \frac{j_J(kr)}{(kr)^2} Y^{TE}_{(JM)ab}(\hat{\mathbf{n}}), \\
\Psi^{k,VE}_{(JM)ab}(\mathbf{x}) =& -\sqrt{3J(J+1)} f_J(kr)\, Y^{L}_{(JM)ab}(\hat{\mathbf{n}}) - \left(j_J(kr) + 2g_J(kr) + 2f_J(kr)\right) Y^{VE}_{(JM)ab}(\hat{\mathbf{n}}) \\
& - \sqrt{(J-1)(J+2)} \left(f_J(kr) + 2\frac{j_J(kr)}{(kr)^2}\right) Y^{TE}_{(JM)ab}(\hat{\mathbf{n}}), \\
\Psi^{k,TE}_{(JM)ab}(\mathbf{x}) =& -\frac{\sqrt{3J(J+1)(J-1)(J+2)}}{2} \frac{j_J(kr)}{(kr)^2} Y^{L}_{(JM)ab}(\hat{\mathbf{n}}) \\
& - \sqrt{(J-1)(J+2)} \left(f_J(kr) + 2\frac{j_J(kr)}{(kr)^2}\right) Y^{VE}_{(JM)ab}(\hat{\mathbf{n}}) \\
& - \frac{1}{2}\left(-j_J(kr) + g_J(kr) + 4f_J(kr) + 6\frac{j_J(kr)}{(kr)^2}\right) Y^{TE}_{(JM)ab}(\hat{\mathbf{n}}), \\
\Psi^{k,VB}_{(JM)ab}(\mathbf{x}) =& -i\left(j'_J(kr) - \frac{j_J(kr)}{kr}\right) Y^{VB}_{(JM)ab}(\hat{\mathbf{n}}) - i\sqrt{(J-1)(J+2)} \frac{j_J(kr)}{kr} Y^{TB}_{(JM)ab}(\hat{\mathbf{n}}), \\
\Psi^{k,TB}_{(JM)ab}(\mathbf{x}) =& -i\sqrt{(J-1)(J+2)} \frac{j_J(kr)}{kr} Y^{VB}_{(JM)ab}(\hat{\mathbf{n}}) - i\left(j'_J(kr) + 2\frac{j_J(kr)}{kr}\right) Y^{TB}_{(JM)ab}(\hat{\mathbf{n}}),
\end{aligned}
\tag{2.61}
$$

where we define radial functions

$$f_J(x) \equiv \frac{d}{dx}\frac{j_J(x)}{x}, \qquad \text{and} \qquad g_J(x) \equiv -j_J(x) - 2f_J(x) + (J-1)(J+2)\frac{j_J(x)}{x^2}. \quad (2.62)$$

For many projected observables on the sky such as CMB temperature and polarization anisotropies, angular power spectra appear as a convolution between the three-dimensional power spectra and these geometrical radial functions.[49]

These results for decomposition are of great practical use. The components proportional to $Y^{TE}_{(JM)ab}$ and $Y^{TB}_{(JM)ab}$ would be of principle interest if only partial components of the entire tensor field is measurable. For instance, ruler distortion of cosmic length scales is generally described by a rank-2 distortion tensor;[56] it may be that the only observable components of the distortion tensor are those transverse to the line of sight.

2.4.3 TAM expansion of tensor plane wave

To make contact between the Fourier expansion and the TAM expansion, it is useful to derive the tensor-field analogue of Rayleigh formula, i.e. a symmetric, trace-free tensor plane wave expanded in terms of tensor TAM basis functions. The result is a straightforward generalization of the case for vector field.

In terms of the OAM basis functions, the expansion is

$$\hat{\varepsilon}_{ab}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}} = \sum_{lJM} 4\pi i^l B^l_{(JM)}(\hat{\mathbf{k}})\Psi^{l,k}_{(JM)ab}(\mathbf{x}), \quad (2.63)$$

Here the plane wave polarization tensor $\hat{\varepsilon}_{ab}(\mathbf{k})$ is normalized to unity. The expansion coeffecients are given by

$$B^{l}_{(JM)}(\hat{\mathbf{k}}) = \hat{\varepsilon}^{ab}(\mathbf{k}) Y^{l\,*}_{(JM)ab}(\hat{\mathbf{k}}). \tag{2.64}$$

Expansions of identical form exist for the LVT basis and the helicity basis

$$\begin{aligned}
\hat{\varepsilon}_{ab}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} &= \sum_{\alpha} \sum_{JM} 4\pi i^{J} B^{\alpha}_{(JM)}(\hat{\mathbf{k}}) \Psi^{k,\alpha}_{(JM)ab}(\mathbf{x}), \\
\hat{\varepsilon}_{ab}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} &= \sum_{\lambda=0,\pm1,\pm2} \sum_{JM} 4\pi i^{J} B^{\lambda}_{(JM)}(\hat{\mathbf{k}}) \Psi^{k,\lambda}_{(JM)ab}(\mathbf{x}),
\end{aligned} \tag{2.65}$$

with expansion coefficients

$$\begin{aligned}
B^{\alpha}_{(JM)}(\hat{\mathbf{k}}) &= \hat{\varepsilon}^{ab}(\mathbf{k}) Y^{\alpha\,*}_{(JM)ab}(\hat{\mathbf{k}}), \\
B^{\lambda}_{(JM)}(\hat{\mathbf{k}}) &= \hat{\varepsilon}^{ab}(\mathbf{k}) Y^{\lambda\,*}_{(JM)ab}(\hat{\mathbf{k}}).
\end{aligned} \tag{2.66}$$

For the LVT basis or the helicity basis, respectively, we have

$$\begin{aligned}
\hat{\varepsilon}_{a}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} &= \sum_{\alpha=L,E,B} \sum_{JM} 4\pi i^{J} A^{\alpha}_{(JM)}(\hat{\mathbf{k}}) \Psi^{k,\alpha}_{(JM)a}(\mathbf{x}), \\
\hat{\varepsilon}_{a}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} &= \sum_{\lambda=-1,0,1} \sum_{JM} 4\pi i^{J} A^{\lambda}_{(JM)}(\hat{\mathbf{k}}) \Psi^{k,\lambda}_{(JM)a}(\mathbf{x}).
\end{aligned} \tag{2.67}$$

The expansion coefficients are given by

$$A^{\alpha}_{(JM)}(\hat{\mathbf{k}}) = \hat{\varepsilon}^{a}(\mathbf{k}) Y^{\alpha\,*}_{(JM)a}(\hat{\mathbf{k}}), \qquad A^{\lambda}_{(JM)}(\hat{\mathbf{k}}) = \hat{\varepsilon}^{a}(\mathbf{k}) Y^{\lambda *}_{(JM)a}(\hat{\mathbf{k}}), \tag{2.68}$$

where we define $Y^{\lambda=\pm2}_{(JM)ab} = 2^{-1/2}\left[Y^{TE}_{(JM)ab} \pm i Y^{TB}_{(JM)ab}\right]$, $Y^{\lambda=\pm1}_{(JM)ab} = 2^{-1/2}\left[Y^{VE}_{(JM)ab} \pm i Y^{VB}_{(JM)ab}\right]$,

and $Y^{\lambda=0}_{(JM)ab} = Y^{L}_{(JM)ab}$. Note that these are closely related to the usual spin-2 spherical harmonics,[47,48]

$$\hat{\varepsilon}^{ab}_{\lambda'}(\mathbf{k}) Y^{\lambda}_{(JM)ab}(\hat{\mathbf{k}}) = {}_{-\lambda}Y_{(JM)}\left(\hat{\mathbf{k}}\right)\delta_{\lambda\lambda'}, \qquad \text{for} \qquad \lambda, \lambda' = 0, \pm 1, \pm 2, \tag{2.69}$$

where $\hat{\varepsilon}^{ab}_{\lambda'}(\mathbf{k})$ is the polarization tensor for a tensor plane wave having wave vector $\mathbf{k}$ and helicity λ. In terms of the polarization vectors defined in Eq. (2.39), they can be constructed as

$$\hat{\varepsilon}^{ab}_{\pm 1} = \frac{1}{\sqrt{2}}\left[\hat{\varepsilon}^{a}_{\pm 1}\hat{n}^{b} + \hat{\varepsilon}^{b}_{\pm 1}\hat{n}^{a}\right], \qquad \hat{\varepsilon}^{ab}_{\pm 2} = -\hat{\varepsilon}^{a}_{\pm 1}\hat{\varepsilon}^{b}_{\pm 1}, \qquad \hat{\varepsilon}^{ab}_{0} = \sqrt{\frac{3}{2}}\left(\frac{1}{3}\delta_{ab} - \hat{n}^{a}\hat{n}^{b}\right), \tag{2.70}$$

2.4.4 Expansion of tensor fields and power spectra

Now that we have constructed a complete set of basis tensor TAM waves, such as the LVT basis, any symmetric, trace-free tensor field admits an expansion in terms of such basis,

$$h_{ab}(\mathbf{x}) = \sum_{JM}\sum_{\alpha}\int \frac{k^2 dk}{(2\pi)^3} h^{\alpha}_{(JM)}(k) 4\pi i^{J} \Psi^{\alpha,k}_{(JM)ab}(\mathbf{x}), \tag{2.71}$$

for $\alpha = L, VE, VB, TE, TB$ with expansion coefficients,

$$h^{\alpha}_{(JM)}(k) = \int d^3\mathbf{x}\, h^{ab}(\mathbf{x})\left[4\pi i^{J}\Psi^{\alpha,k}_{(JM)ab}(\mathbf{x})\right]^{*}. \tag{2.72}$$

Similar relations hold for the OAM basis or the helicity basis. If a stochastic tensor field is statistically homogeneous *and* isotropic, the VE and VB modes must share the same

power spectrum; likewise, the TE and TB modes must share the same power spectrum. Otherwise, statistical homogeneity will be violated. On the other hand, those may differ from the power spectrum for the L mode. Therefore, there exist three independent power spectra for TAM waves of different polarization type,

$$\left\langle \left[h^{\alpha}_{(JM)}(k)\right]^{*} h^{\beta}_{(J'M')}(k')\right\rangle = P_T(k)\delta_{JJ'}\delta_{MM'}\delta_{\alpha\beta}\frac{(2\pi)^3}{k^2}\delta_D(k-k'), \qquad \alpha,\beta = TE, TB. \tag{2.73}$$

$$\left\langle \left[h^{\alpha}_{(JM)}(k)\right]^{*} h^{\beta}_{(J'M')}(k')\right\rangle = P_V(k)\delta_{JJ'}\delta_{MM'}\delta_{\alpha\beta}\frac{(2\pi)^3}{k^2}\delta_D(k-k'), \qquad \alpha,\beta = VE, VB. \tag{2.74}$$

$$\left\langle \left[h^{L}_{(JM)}(k)\right]^{*} h^{L}_{(J'M')}(k')\right\rangle = P_L(k)\delta_{JJ'}\delta_{MM'}\delta_{\alpha\beta}\frac{(2\pi)^3}{k^2}\delta_D(k-k'). \tag{2.75}$$

If, on the other hand, there exists non-vanishing cross-correlation between VE and VB modes, or between TE and TB modes, parity conservation must be statistically broken (see Ref.[51] for an application to CMB polarization anisotropies).

2.5 Summary

The TAM formalism may be utilized to solve many problems in cosmology. It is a powerful calculational tool for quantitative study of cosmological perturbations, both at the linear order and possibly extendable to second order in perturbations,[44] and especially when vector or tensor fields are involved in the problem. Applications of the TAM formalism include all-sky analysis of weak gravitational lensing by general metric perturbations,[42,57,58] relic bispectra in LSS from tensor perturbations,[59–61] CMB polarization power spectra,[51]

fossil bispectra in the CMB,[44,62] and all-sky dark matter directional detection.[63]

Without presenting explicit examples here (as the formalism will be used in later chapters in any case), we briefly summarize the standard procedure for utilizing the TAM formalism. One first determines the spatial rotation representation to which the perturbation field under investigation belongs, namely, the spin of the perturbation field and its polarization type. For the latter, it is often convenient to work with the longitudinal/transverse basis for vector TAM waves, or the LVT basis for tensor TAM waves (although in special cases, where parity violating effects are under scrutiny, it may be more advantageous to use the helicity basis). For instance, matter density field should be expanded in the scalar TAM basis. Moreover, the infall velocity field from gravitational clustering is purely a potential flow, and hence should be expanded using the longitudinal vector TAM wave $\Psi^{k,L}_{(JM)a}(\mathbf{x})$; on the other hand, velocity vorticity, if under discussion, must be expanded using the transverse vector TAM waves $\Psi^{k,E}_{(JM)a}(\mathbf{x})$ and $\Psi^{k,B}_{(JM)b}(\mathbf{x})$. Likewise, gravitational wave should be expanded using the transverse, trace-free tensor TAM waves $\Psi^{k,TE}_{(JM)ab}(\mathbf{x})$ and $\Psi^{k,TB}_{(JM)ab}(\mathbf{x})$.

One then derives the linearized equation of motion for each TAM mode much like what is conventionally done for a Fourier mode. If dynamics has translational and rotational symmetry, which follow from spatial homogeneity and isotropy, the equation of motion for each TAM wave is independent, and only depends on the wave number k and the possibly the polarization type, but not on the total angular momentum J and M. This implies that the TAM basis and the Fourier basis share the same sets of equation of motion, or transfer functions, for linearized perturbations. The TAM formalism is advantageous because a spherical harmonic observable on the sky labelled by J, M receives contribution *only* from

TAM modes with the same value for the total angular momentum. As a result, a further convolution unavoidable in the Fourier approach is skipped, as the geometrical projection factors have been built into the construction of the TAM basis functions.

For vector fields, parity-even observables only receive contribution from $\alpha = L, E$ polarization modes, while parity-odd observables are dependent on $\alpha = B$ polarization mode. Similarly, for symmetric, trace-free tensor fields, parity-even observables only receive constribution from $\alpha = L, VE, TE$ modes, while parity-odd observables are dependent on $\alpha = VB, TB$ modes. A good example is the CMB polarization anisotropy, part of which may be generated by a stochastic background of the primordial gravitational waves. In this example, the TE components of the gravitational wave induce the E-mode polarization, and the TB components induce the B-mode polarization.[51]

Finally, we would like to end this chapter with remarks on the extension of TAM formalism to CMB radiation transfer. Unlike metric perturbations, or perturbations in matter energy stress tensor, the radiation transfer problem deals with perturbations to the phase space distribution, which is not only a function of spatial position $\mathbf{x}$, but also the photon direction $\hat{\mathbf{n}}$ (frequency dependence is often approximated by an equilibrium Planck blackbody distribution unless spectrum distortion needs to be considered). Perturbations in phase space density is often decomposed into temperature multipoles Θ_ℓ — monopole ($\ell = 0$), dipole ($\ell = 1$), quadrupole ($\ell = 2$), octupole ($\ell = 3$) and so on — according to an increasingly higher-order direction dependence, and similar generalization for polarization. Following the TAM formalism, the monopole behaves as a scalar field, the dipole as a spin-1 vector field, the quadrupole as a spin-2 tensor field, and in general the ℓth-multipole as a

spin-ℓ field. An extension of Ref.[44] to fields of higher spin is therefore desired.

There has been attempt to construct spherical basis to expand radiation phase space perturbations,[49] although in that reference the orbital angular momentum from the position dependence part (a plane-wave dependence) and the "spin" from the direction dependence part are still decoupled, and therefore do not quite follow the philosophy of TAM formalism.

Chapter 3

Inflationary Fossil in CMB

The idea of inflation provides a causal explanation for initial cosmological perturbations that appear to have correlation on superhorizon scales. Single-field slow-roll inflation makes the prediction that the primordial curvature perturbation $\zeta(\mathbf{x})$, as defined through

$$ds^2 = a^2(\tau)\left[-d\tau^2 + e^{2\zeta(\mathbf{x})}\delta_{ij}dx^i dx^j\right] \tag{3.1}$$

when all perturbations of interests have wavelengths much longer than the horizon scale, has a Gaussian random distribution, described by a nearly scale-invariant power spectrum[20]

$$\left\langle \zeta(\mathbf{k})\zeta(\mathbf{k}')\right\rangle \;=\; (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')A_s\left(\frac{k}{k_0}\right)^{n_s-4}. \tag{3.2}$$

where $\ln(10^{10}\,A_s) \approx 3.098$ and $n_s \approx 0.9619$,[64] and k_0 is a chosen pivot scale. This sets up an initial adiabatic perturbation for CMB fluctuations and large-scale structure formation,

which means that different Hubble patches in the early Universe have mis-synchronized age and therefore advance or delay with respect to one another along a *unique* cosmographic and thermal history. The mis-synchronization, usually termed *single-clock inflation*, can be characterized by an initial gravitational potential in the conformal Newtonian gauge, $\Phi(\mathbf{x}) = -(3/5)\zeta(\mathbf{x})$ during matter domination. Since the subsequent evolution on subhorizon scales is linear before perturbations grow large, this predicts Gaussian random fluctuation in the CMB, as well as Gaussian random density fluctuation of Cold Dark Matter.

Nevertheless, a nearly scale-invariant, Gaussian power spectrum for primordial curvature perturbation is a very generic prediction of single-field inflation with a slow-roll attractor solution.[17,20] In this minimal scenario, an adiabatic scalar mode with additional tensor metric modes (gravitational waves) are the only physical degrees of freedom. While the phenomenology is universal, and therefore has much predictive power, it reveals little about the physical origin of the adiabatic mode, apart from the existence of a slow-roll phase and a quasi-de Sitter expansion.

It is conceivable that the inflationary physics is far richer than what is characterized in the single-field slow-roll paradigm. In particular, it is not reasonable to take it for granted that a single scalar field is the only physical degree of freedom (apart from gravitational waves) relevant around the energy scale of inflation. In many inflation models, extra particle fields are introduced to address various problems from elementary particle physics considerations, such as naturalness, UV completion, or supersymmetry.[65] New scalar fields, elementary or composite, might play a role in extensions of single-field inflation.[66–70] The adiabatic scalar mode might represent a dominant slow-roll direction in a

configuration space of multiple fields.[71] Spin-one gauge fields might have driven inflation and have induced departure from statistical isotropy.[72–74] If general relativity is modified at high energies, new vector degrees of freedom in the metric may arise.[75] Even in the simplest case of general relativity, tensor metric perturbations can be dynamically important.[76–78] All those alternative scenarios predict complex statistics for the primordial adiabatic mode that differ from the nearly Gaussian one for single-field slow-roll inflation, a powerful diagnostic of inflationary dynamics in the distant past through a survey of inhomogeneities in the late-time Universe. The leading-order departure is a three-point correlation function in real space, or a bispectrum in Fourier space, which would (nearly) vanish for single-field slow-roll.

We devote this Chapter to a general study of primordial bispectra, or non-Gaussianity, as a hopeful probe for new inflationary physics. We examine multi-spectra that cross-correlate Fourier modes of the primordial curvature perturbation (or equivalently the initial gravitational potential) and Fourier modes of an extra field. Here, the extra field will be referred to as a "fossil" field, in the sense that no direct cosmological tracer for it is available in the late-time Universe, unlike CMB anisotropies and LSS as tracers for the primordial curvature perturbation. Indeed, no cosmological or astrophysical tracer for any new field in the primordial universe is observationally known; on the other hand, it is likely that physical degrees of freedom at inflation energy scale decouple at low energies from the Standard Model of elementary particles, or even from the Dark Matter. Nevertheless, lack of direct tracers does not always imply observational irrelevance. "Fossil" fields may interact significantly with the primordial curvature perturbation during inflation and may alter the

statistical property of the curvature perturbation.

More specifically, one Fourier mode of the "fossil" field may cross-correlate with curvature perturbations in a primordial scalar-scalar-fossil bispectrum (panel (a) of Figure 3.1), signifying a modulation of curvature two-point correlation by the extra field. This induces an apparent breakdown of statistical isotropy in the curvature perturbation when the fossil field is not observationally accessible. This is equivalent to the existence of a *connected* trispectrum — four-point correlation in Fourier space — for the curvature perturbation but mediated by an internal "fossil" mode. In the presence of a stochastically generated background of "fossil" field. The trispectrum represents ensemble-averaged effects of the fossil bispectrum depicted in (a) of Figure 3.1.

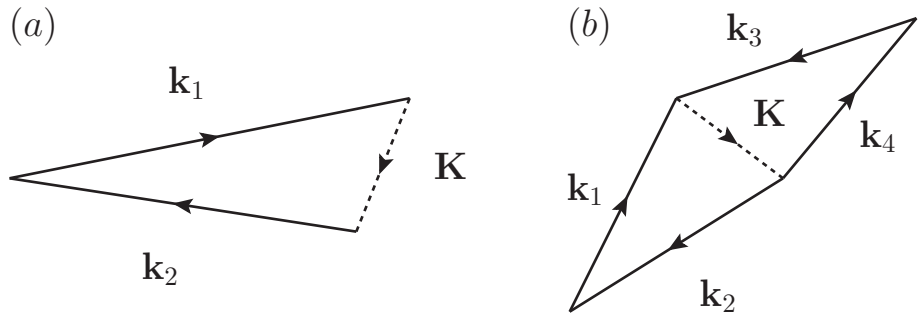


Figure 3.1: (a) A (cross-)bispectrum involving two Fourier modes of the curvature perturbation and a third Fourier mode of a "fossil" field. (b) A trispectrum for the curvature perturbation mediated by an internal mode of "fossil" field. Solid lines represent the curvature perturbation, and dashed lines represent the "fossil" field.

This Chapter is mainly based on the published work of Ref.[62] We will start with a discussion of the primordial scalar-scalar-fossil bispectrum in Section 3.1. We briefly review the fossil estimator for large-scale structure surveys. We then turn to Section 3.2 for a detailed discussion of detecting its signature in the CMB using a bipolar spherical

harmonic parametrization. Finally, concluding remarks on this subject will be given in Section 3.3.

Starting from this Chapter, we switch to lower-case latin letters $i, j, k, \cdots$ for spatial indices in three dimensions, slightly different from the notation adopted in Chapter 2.

3.1 Fossil Bispectrum

To derive from first principle the primordial scalar-scalar-fossil bispectrum, one generally would need to first derive the cubic interaction terms from the action describing the inflationary physics, and then adopt the *in-in formalism* of quantum field theory to compute the three-point function toward the end of inflation.[79] The exact result of the calculation is dependent on the specific inflation model. From a practical point of view, since we do not have the knowledge about the details of inflationary physics, a *model-independent* parametrization will be pursued here.

Following the discussion of Ref.,[43] let us consider two Fourier modes of primordial potential perturbation [1], having wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$, respectively. Similarly, consider a plane wave of generic fossil field h, having wave vector $\mathbf{K}$ and polarization type p (in case the fossil field carries non-zero spin). The most general form for the scalar-scalar-fossil

[1] In order to conveniently connect to late-time observables, we will directly write down the scalar-scalar-fossil bispectrum in terms of the gravitational potential during matter domination Φ, as opposed to the curvature perturbation ζ, despite that the latter is a more fundamental perturbation variable for adiabatic initial condition.

bispectrum reads

$$\langle \Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)h^\alpha(\mathbf{K})\rangle = (2\pi)^3\delta_D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{K})\, f_h^\alpha(k_1,k_2,K)\, P_h^\alpha(K)\, \epsilon_{ij}^\alpha(\mathbf{K})\hat{k}_1^i\hat{k}_2^j. \qquad (3.3)$$

The Dirac-δ function, a consequence of translational invariance of the Universe, imposes that a non-zero bispectrum only exists if the three wave vectors sum up to zero. The bispectrum can be factorized into the product of the (isotropic) fossil power spectrum $P_h^\alpha(K)$ for polarization type α, and a bispectrum shape function $f_h^\alpha(k_1,k_2,K)$ dependent on the three wave numbers as well as (possibly) the polarization type. The last tensorial structure $\epsilon_{ij}^\alpha\hat{k}_1^i\hat{k}_2^j$, where $\hat{k}_1^i$ and $\hat{k}_2^j$ being the wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ normalized to unity, follows from the fact that the bispectrum might be linearly proportional to the fossil *polarization tensor* $\epsilon_{ij}^\alpha(\mathbf{K})$, for the reason that superposition of two different polarizaton states is also a possible polarization state.

A note of clarification here. To keep our discussion general, we take into account the possibilities that the fossil field may be a scalar field, a vector field, or a spin-2 tensor field. Physical degrees of freedom with spin higher than two, though not proven impossible in the primordial Universe, go beyond the scope of our discussion. Though not necessary, we choose to use a rank-two symmetric polarization tensor $\epsilon_{ij}^\alpha(\mathbf{K})$ to describe all three cases of spin, for uniformality of our analysis. A scalar field would correspond to a polarization tensor $\epsilon_{ij}^\alpha(\mathbf{K})$ being a pure trace $\propto \delta_{ij}$, or of the longitudinal type $\propto (\hat{K}_i\hat{K}_j - \delta_{ij}/3)$. To represent vector fields, the polarization tensor takes the transverse-vectorial form $\propto \hat{K}_{(i}w_{j)}$ with w_i being a perpendicular, unit vector $w_iK^i = 0$. Finally, a tensor fossil field would have

its polarization completely perpendicular to the direction of propagaton, i.e. $K^i\epsilon^\alpha_{ij}(\mathbf{K}) = 0$. With our formalism, all three cases of spin can be uniformly specified by a polarization type α.

The tensorial structure $\epsilon^\alpha_{ij}\hat{k}^i_1\hat{k}^j_2$ fixes the polarization dependence of the bispectrum. If one choose the direction of $\mathbf{K}$ to be along the z axis, and with respect to this axis one characterizes the orientation of the Fourier-space triangle, i.e. the closed triangle formed by the three wave vectors, by an azimuthal angle φ, as shown in Figure 3.2, then a bispectrum independent of φ is expected from a scalar fossil field. A dependence $\sim e^{\pm i\varphi}$ is an indication that the fossil field is a spin-one vector. Likewise, a dependence $\sim e^{\pm 2i\varphi}$ would arise from a spin-two tensor fossil field.

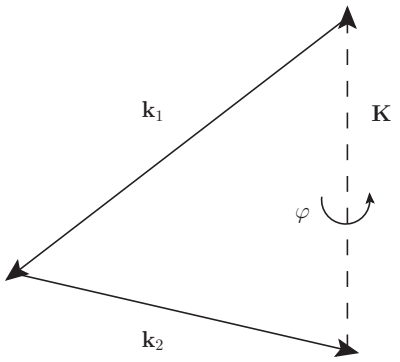


Figure 3.2: How the scalar-scalar-fossil bispectrum is dependent on the azimuthal angle φ about the fossil mode reveals the spin of the fossil field.

Eq. (3.3) implies a modulation of the observable scalar potential two-point statsitics by the fossil field that has become dynamically irrelevant in the post-inflationary Universe. An equivalent mathematical expression is the scalar two-point correlation *in the presence*

of a fossil mode,

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\rangle_{h^\alpha(\mathbf{K})} = (2\pi)^3\delta_D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{K})f_h^p(k_1,k_2,K)\,\epsilon_{ij}^\alpha(\mathbf{K})\hat{k}_1^i\hat{k}_2^j[h^\alpha(\mathbf{K})]^*. \tag{3.4}$$

3.1.1 Fossil estimator for large-scale structure

How would one go about measuring this correlation? First of all, each triangular configuration provides an estimator for the fossil field realization

$$\widehat{h_\alpha(\mathbf{K})} = \Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\left[f_h^\alpha(k_1,k_2,K)\epsilon_{ij}^\alpha(\mathbf{K})\hat{k}_1^i\hat{k}_2^j\right]^{-1}. \tag{3.5}$$

All triangular configurations with the same $\mathbf{K}$ can be weighted with the inverse variance to form a *minimum-variance estimator* for $h_\alpha(\mathbf{K})$. This would be a true minimum-variance estimator if the estimator for each triangle is Gaussian. Actually, if the potential field Φ is a Gaussian random field from the initial condition, the estimator $\widehat{h_\alpha(\mathbf{K})}$ satisfies the χ^2-distribution with two independent degrees of freedom.[80] However, inverse-variance weighting is still a good approximation for an effective reduction of sample variance. We have the (nearly) minimum-variance estimator

$$\widehat{h_\alpha(\mathbf{K})} = V^{-1}P_\alpha^n(K)\sum_{\mathbf{k}}\frac{f_\alpha(k,|\mathbf{K}-\mathbf{k}|,K)\epsilon_{ij}^\alpha(\mathbf{K})k^i(K-k)^j}{2k|\mathbf{K}-\mathbf{k}|P_\Phi^{\rm tot}(k)P_\Phi^{\rm tot}(|\mathbf{K}-\mathbf{k}|)}\Phi(\mathbf{k})\Phi(\mathbf{K}-\mathbf{k}), \tag{3.6}$$

where V is the total survey volume (it is introduced for normalization of Fourier modes, and will cancel in the final result), and $P_\Phi^{\rm tot}(k) = P_\Phi(k) + P_\Phi^n(k)$ is the potential power

spectrum including possible shot noise. The *noise power spectrum* $P_\alpha^n(K)$ is given by

$$P_\alpha^n(K) = \left[V^{-1} \sum_{\mathbf{k}} \frac{\left| f_\alpha(k, |\mathbf{K}-\mathbf{k}|, K) \epsilon_{ij}^\alpha(\mathbf{K}) k^i (K-k)^j \right|^2}{2k^2 |\mathbf{K}-\mathbf{k}|^2 P_\Phi^{\rm tot}(k) P_\Phi^{\rm tot}(|\mathbf{K}-\mathbf{k}|)} \right]^{-1}. \tag{3.7}$$

Due to statistical isotropy, $P_\alpha^n(K)$ is expected to be isotropic, i.e. only dependent on the magnitude of $\mathbf{K}$ but not its direction.

If the fossil field background is also nearly Gaussian, it may be parametrized by an (isotropic) power spectrum

$$\left\langle h_\alpha(\mathbf{K}) h_\alpha(\mathbf{K'}) \right\rangle = (2\pi)^3 \delta_D(\mathbf{K}+\mathbf{K'})\, A_h^\alpha P_h^\alpha(K), \tag{3.8}$$

with a fiducial shape $P_h^\alpha(K)$ and an unknown constant amplitude A_h^α. An estimator for A_h^α may be written down for each fossil Fourier mode $\mathbf{K}$,

$$\widehat{A_h^\alpha(\mathbf{K})} = [P_h^\alpha(K)]^{-1} \left[V^{-1} \left| \widehat{h_\alpha(\mathbf{K})} \right|^2 - P_\alpha^n(K) \right]. \tag{3.9}$$

Since each estimator $\widehat{h_\alpha(\mathbf{K})}$ has been built from a large number of triangular configurations in Fourier space, its probability distribution is close to Gaussian due to the Central Limit Theorem. Therefore, we may once again adopt inverse-variance weighting to construct a minimum-variance estimator for the amplitude A_h^α,

$$\widehat{A_h^\alpha} = (\sigma_h^\alpha)^2 \sum_{\mathbf{K},\alpha} \frac{[P_h^\alpha(K)]^2}{2\,[P_\alpha^n(K)]^2} \left(V^{-1} \left| \widehat{h_\alpha(\mathbf{K})} \right|^2 - P_\alpha^n(K) \right), \tag{3.10}$$

with

$$(\sigma_h^\alpha)^{-2} = \sum_{\mathbf{K},\alpha} \frac{[P_h^\alpha(K)]^2}{2\,[P_\alpha^n(K)]^2}. \tag{3.11}$$

The summation over polarization in Eq. (3.10) requires explanatory remarks: For a scalar fossil, there is only a single, unique polarization state, and summation over α is not needed. If the fossil is a transverse vector field, one sums over the two transverse vector polarizations that are related to each other by $\pm 90^\circ$ rotations about $\mathbf{K}$. If the fossil is a transverse tensor field, one sums over the two transverse tensor polarizations that are related to each other by $\pm 45^\circ$ rotations about $\mathbf{K}$. Note that if statistical isotropy is preserved, fossil power spectrum $A_h^\alpha P_h^\alpha(K)$ must be the same between the two transverse vector polarizations, and between the two transverse tensor polarizations as well.

For given survey of primoridial perturbations, the survey volume determines the smallest Fourier wavenumber $K_{\rm min}$ that can be probed, and the survey resolution determines the largest Fourier wavenumber $K_{\rm max}$ resolvable. Assuming all Fourier modes of linear potential perturbation to be statistically independent of each other, a larger *dynamical ratio* $K_{\rm max}/K_{\rm min}$ would then imply a larger amount of statistically independent information, and a larger number of triangular configurations with which one measures the bispectrum. Therefore, for given fiducial $P_h^\alpha(K)$ and fiducial $f_\alpha(k_1, k_2, k_3)$, one can test against the null hypothesis that the potential $\Phi(\mathbf{k})$ is Gaussian. The criterion would be that the measured optimal estimator $\widehat{A_h^\alpha}$ has high significance ($\gtrsim 3\sigma$) over the variance expected from a Gaussian random $\Phi(\mathbf{k})$.

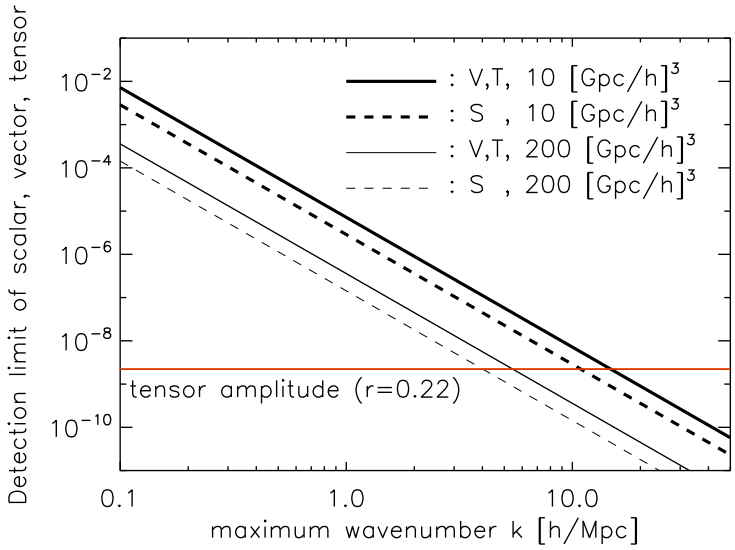


Figure 3.3: (From Ref.[43]) Sensitivity to the fossil amplitude A_h^p as dependent on survey volume and resolution.

As an illustrating example, Ref.[43] studies a scale-invariant fossil power spectrum $P_h^\alpha(K) = K^{-3}$ [2]. The authors assume that the scalar-scalar-fossil bispectrum is dominated by squeezed configurations $K \ll k_1, k_2$ with order-unity non-Gaussianity $f_h^\alpha(k_1, k_2, K) \approx -(3/2)P_\Phi(k_1)$. Figure 3.3 presents the detection sensitivity to the fossil amplitude A_h^α as a function of $K_{\rm max}$ and $K_{\rm min}$. The scalar-scalar-tensor bispectrum in single-field inflation is chosen as a benchmark signal to target. In this case, the significance improves as $(K_{\rm max}/K_{\rm min})^3$. While current or upcoming galaxy surveys fall short in the dynamical ratio $K_{\rm max}/K_{\rm min} \lesssim 100$, future high-redshift 21cm surveys probing a comoving volume ~ 100 [Gpc/h]^3 and with a resolution up to a few h/Mpc in comoving wave number (which corresponds to $K_{\rm max}/K_{\rm min} \sim 5000$), should in principle be capable of detecting squeezed fossil non-Gaussianity down to $A_h^\alpha \sim 10^{-8}$. This translates into a favorable fossil field fluc-

[2]Note that instead of measuring potential-potential-fossil correlation, the authors of Ref.[43] choose to measure density-density-fossil correlation. However, in the ideal case of zero survey noise, the conversion between potential and density evidently drops out from Eq. (3.6) and Eq. (3.7).

tuation at the same order of magnitude $\sim 10^{-5}$ as that for the primordial potential. The constraining power can be a factor of few better in alternative theories of inflation.[81]

3.2 Fossil Bispectrum in CMB

In Section 3.1.1, we have briefly reviewed the method to detect a primordial scalar-scalar-fossil bispectrum from tracers of matter density distribution, including galaxies and the cosmological 21cm emission. With a redshift survey, these trace the matter density field, and therefore infer the initial potential perturbation, within a three dimensional survey field.

There is another powerful cosmological probe of initial perturbations — the fluctuations in the CMB. Unlike redshift surveys of LSS, the CMB is a two dimensional map from the last scattering surface when electrons and protons recombine. Although the information content of a two dimensional survey improves less rapidly with the resolution as for a three dimensional survey, CMB anisotropies are unbiased tracers of the initial perturbation and are less undermined by observational systematics. Latest CMB experiments, such as *Planck*, has been able to measure temperature and polarization anisotropies down to the cosmic variance limitation over a wide range of angular scales $2 \leq l \lesssim 1500$.[82] This large dynamical ratio partly compensates for the two dimensional nature of the observation, making CMB still the most effective avenue so far to constrain primordial non-Gaussianity.[83]

In this Section, we develop the basic formalism to constrain a primordial scalar-scalar-fossil bispectrum using CMB observations. For full-sky CMB measurement performed on the spherical sky, spherical harmonic analysis is the conventional treatment. Therefore, we

choose to parameterize perturbations using the TAM formalism developed in Chapter 2 throughout our calculation. To this end, Eq. (3.4) can be rewritten in terms of the TAM modes

$$\langle \Phi_{l_1 m_1}(k_1)\Phi_{l_2 m_2}(k_2)\rangle_{h^\alpha_{JM}(K)} = [h^\alpha_{JM}(K)]^* f^\alpha_h(k_1,k_2,K)$$
$$\times (4\pi)^3(-i)^{l_1+l_2+J}\frac{1}{k_1k_2}\int d^3\mathbf{x}\left(\nabla^i\Psi^{k_1}_{(l_1m_1)}(\mathbf{x})\right)\left(\nabla^j\Psi^{k_2}_{(l_2m_2)}(\mathbf{x})\right)\Psi^{\alpha,K}_{(JM)ijs}(\mathbf{x}). \qquad (3.12)$$

This simply says that there may exist non-zero correlation between two different TAM modes of the potential, $\Phi_{l_1m_1}(k_1)$ and $\Phi_{l_2m_2}(k_2)$, if a fossil TAM mode $h^\alpha_{JM}(K)$ (which will not be directly observable) is present. Note that in this form the overlap of three TAM basis wavefunctions is involved. Explicit expressions for these three-mode overlaps are presented in Ref.[44] Depending on the spin of the fossil field, appropriate polarization type α should be chosen. For a scalar fossil field, one chooses the longitudinal tensor TAM wave $\alpha = L$. A vector fossil field is associated with the two transverse vectorial TAM modes $\alpha = VE, VB$. The two transverse tensorial TAM modes $\alpha = TE, TB$ are related to a tensor fossil field. A statistically homogeneous and isotropic scalar-scalar-fossil interaction would impose the relations $f^{VE}_h(k_1,k_2,K) = f^{VB}_h(k_1,k_2,K) \equiv f^{V}_h(k_1,k_2,K)$ and $f^{TE}_h(k_1,k_2,K) = f^{TB}_h(k_1,k_2,K) \equiv f^{T}_h(k_1,k_2,K)$.

3.2.1 Bipolar Spherical Harmonics

The correlation functions for temperature and polarization anisotropies are isotropic in the absence of a scalar-scalar-fossil bispectrum. Taking the temperature for example, the

covariance matrix for multipole moments is diagonal

$$\left\langle a^T_{l_1 m_1} a^T_{l_2 m_2} \right\rangle = C^{TT}_{l_1} \delta_{l_1 l_2} \delta_{m_1 m_2}, \tag{3.13}$$

where $\langle \cdots \rangle$ represents ensemble average. Transforming back into real space, the two-point correlation function only depends on the angular separation. However, when a *given realization of the fossil field* correlates with the initial potential perturbation, statistical isotropy is explicitly broken. The most general covariant matrix is shown to have the form[84] ($\langle \cdots \rangle_h$ denotes a conditional ensemble average for fixed fossil configuration)

$$\left\langle a^T_{l_1 m_1} a^{T*}_{l_2 m_2} \right\rangle_h = C^{TT}_{l_1} \delta_{l_1 l_2} \delta_{m_1 m_2} + \sum_{JM} (-1)^{m_2} \langle l_1 m_1 l_2, -m_2 | JM \rangle A^{JM}_{l_1 l_2}, \tag{3.14}$$

where we have introduced the Clebsch-Gordan coefficients with the conventional symbol $\langle l_1 m_1 l_2 m_2 | l_3 m_3 \rangle$. The second term, signaling a violation of statistical isotropy, is parametrized through the bipolar spherical harmonic (BiPoSH) coefficients,

$$A^{JM}_{l_1 l_2} = (-1)^{l_1 + l_2 + M} \sqrt{2J+1} \sum_{m_1 m_2} (-1)^{m_2} \begin{pmatrix} l_1 & l_2 & J \\ m_1 & -m_2 & -M \end{pmatrix} \left\langle a^T_{l_1 m_1} a^{T*}_{l_2 m_2} \right\rangle_h . \tag{3.15}$$

The array in round brackets stands for the standard Wigner-3j symbol, which is essentially a more symmetric form for the Clebsch-Gordan coefficients. The Wigner-3j symbol imposes that BiPoSH components exist only for $|l_1 - l_2| \leq J \leq l_1 + l_2$. On the other hand, we restrict our analysis to quadrupolar modulation or higher $J \geq 2$, which can be induced by scalar, vector and tensor fields. In the absence of a scalar-scalar-fossil bispectrum, only the

$J = 0$, $M = 0$ and $l_1 = l_2$ BiPoSH component is non-zero, reducing to the usual angular power spectrum $C_{l_1}^{TT}$.

There are two categories of BiPoSH coefficients, according to how they transform under a flip of parity:[85–87] parity-even BiPoSHs have $J + l_1 + l_2$ equal to an even integer, and parity-odd BiPoSHs have $J + l_1 + l_2$ evaluated to an odd integer. These two categories have the opposite parity under a spatial reflection.

Expanding the fossil field by the TAM basis is particularly advantageous for making contact with the BiPoSH formalism. BiPoSH components with given total angular momentum quantum numbers J, M can be induced through fossil bispectrum by only those TAM components of the fossil field having the same total angular momentum J, M; the wave number k, on the other hand, has to be integrated over. Moreover, parity-even BiPoSHs can only be generated by TAM waves with the same parity, i.e. $\alpha = L$ for scalar fossil, $\alpha = VE$ for vector fossil, or $\alpha = TE$ for tensor fossil. Likewise, parity-odd BiPoSHs are only generated by odd-parity TAM waves, i.e. $\alpha = VB$ for vector fossil, or $\alpha = TB$ for tensor fossil; a scalar fossil field does not generate any partiy-odd BiPoSH coefficient.

When there exists a hierarchy of angular scales $J \ll l_1, l_2$, the interpretation for the BiPoSH coefficients as a signature for departure from statistical isotropy becomes simple and intuitive — they describe power of fluctuation on small angular scales being spatially modulated by large-angular-scale fluctuations. For circular hot/cold spots of a typical angular scale, fossil modulation systematically distorts their sizes and shapes. Recall that a three dimensional TAM wave $\Psi^{k,\alpha}_{(JM)ij}(\mathbf{x})$ may be expanded into tensor spherical har-

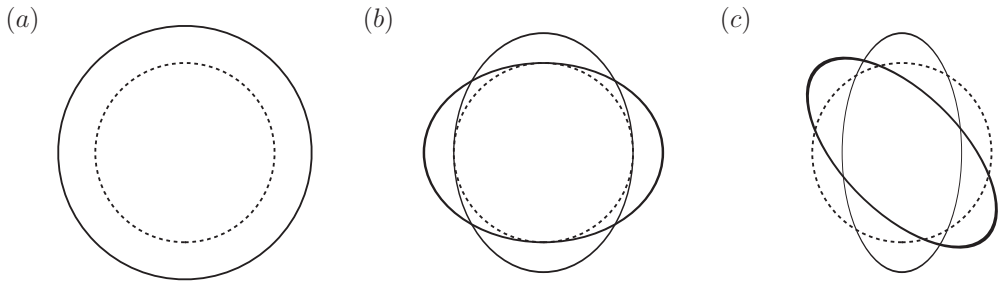


Figure 3.4: Three types of local distortion to a circular hot/cold spot due to modulation by large-angular-scale fossil field: (a) scalar type isotropic (de)-magnification. (b) vectorial type elongation along one of the two transverse directions. (c) tensorial type shear along one of the two choices of major axis separated by 45°. The dashed circle is the original isothermal contour, and solid ones represent distorted contours. For illustrative purpose, the level of distortion has been greatly exaggerated.

monics $Y^{\alpha}_{(JM)ij}(\hat{\mathbf{n}})$. When viewed on the two dimensional last scattering surface from a projected perspective, as shown in Figure 3.4, three types of local distortions may arise: (1) Scalar components proportional to $Y^{L}_{(JM)ij}(\hat{n})$ induce (de-)magnification of the size of the hot/cold spots. (2) Vectorial components proportional to $Y^{VE,VB}_{(JM)ij}(\hat{\mathbf{n}})$ induce elongation of the isothermal contour along a preferred direction. (3) Tensorial components proportional to $Y^{TE,TB}_{(JM)ij}(\hat{\mathbf{n}})$ induce shear of the isothermal contour aligned with a preferred set of perpendicular major axis. We would like to emphasize the point that for a given TAM mode, the foregoing three types of distortion are usually simultaneously present. Due to a reduced amount of information that can be extracted from a two-dimensional projected survey, it is anticipated that scalar, vector and tensor fossil fields cannot be geometrically distinguished, unlike the case of a three-dimensional redshift survey of LSS.[43]

The temperature multipole moments are related to the TAM components of the pri-

mordial potential perturbation through[88]

$$a_{lm}^{T} = \frac{1}{2\pi^2}(-i)^l \int k^2 dk\, g_l^T(k)\Phi_{lm}(k), \tag{3.16}$$

where $g_l^T(k)$ is the radiation transfer function for temperature fluctuation. The BiPoSH coefficients can be straightforwardly derived from Eq. (3.12), Eq. (3.15) and Eq. (3.16). The result reads

$$\begin{aligned} A_{l_1 l_2}^{JM}|_{h_{JM}^{\alpha}(K)} &= -(-i)^J(-1)^{l_1+l_2+P(\alpha)} h_{JM}^{\alpha}(K)\frac{16}{\pi}\left(\frac{(2l_1+1)(2l_2+1)}{4\pi}\right)^{\frac{1}{2}} \\ &\times \int k_1^2 dk_1\, g_{l_1}^T(k_1)\int k_2^2 dk_2\, g_{l_2}^T(k_2)\, f_h^{\alpha}(k_1,k_2,K)\, \mathcal{I}_{l_1 l_2 J}^{\alpha}(k_1,k_2,K), \end{aligned} \tag{3.17}$$

The parity $P(\alpha)$ equals zero for $\alpha = L, VE, TE$ and equals unity for $\alpha = VB, TB$. We have also introduced the overlap of three TAM wavefunctions

$$\begin{aligned} \mathcal{I}_{l_1 l_2 J}^{\alpha}(k_1,k_2,K) &= \left[\frac{4\pi}{(2l_1+1)(2l_2+1)(2J+1)}\right]^{1/2} \\ &\times \left[\int d^3\mathbf{x}\, \Psi_{(l_1 m_1)}^{L,k_1,a}(\mathbf{x})\Psi_{(l_2 m_2)}^{L,k_2,a}(\mathbf{x})\Psi_{(JM)ab}^{\alpha,K}(\mathbf{x})\right]\begin{pmatrix} J & l_2 & l_1 \\ M & m_2 & m_1 \end{pmatrix}^{-1}, \end{aligned} \tag{3.18}$$

which accounts for the non-trivial geometry of a spherical-wave expansion. It can be straightforwardly evaluated using the result of Ref.[44]

In reality, the Gaussian, statistically isotropic CMB sky is found to be distorted due to gravitational deflection of line of sight by intervening large scale matter distribution between the last scattering surface and us.[89–92] The distortion correlates with the temperature

fluctuations on large angular scales, because the same large-scale matter distribution at low redshifts (during Dark Energy domination) would redshift/blueshift CMB photons via the integrated Sachs-Wolfe (ISW) effect. This weak gravitational lensing effect constitutes an irreducible background, i.e. the lensing-ISW bispectrum,[90] in the search of primordial non-Gaussianity.

Neverthessless, lensing signatures are distinct from fossil modulation, if the fossil field has a red power spectrum (e.g. a scale invariant one) and the scalar-scalar-fossil bispectrum has a local shape. In the concordance cosmology, lensing by large-scale gravitational potential is peaked at intermediate angular scales $J \sim 60$. By comparison, effects from a fossil field is primarily present on very large modulation scales $J \lesssim 5$, as our numerical study has indicated. Therefore, the dependence on angular scales can serve as a powerful diagnostic to differenciate between mundane gravitational weak lensing and more exotic cause of non-Gaussianity in the primordial Universe.

Besides, lensing converts a fraction of the E-mode polarization into B-mode polarization on small angular scales $l \gtrsim 200$, which has been measured by ground-based high-resolution deep surveys.[93,94] By comparison, primordial fossil fields do not generate B-mode polarization. It is then conceivable that a comparison between the E-mode and B-mode polarization fields on relevant angular scales would allow for reconstruction of the gravitational lensing potential, which is expected to fully correlate with the lensing-ISW non-Gaussianity but be unrelated to any primordial fossil field.

Finally, since lensing by foreground large-scale structure is generated by a scalar grav-

itational potential, the lensing-ISW non-Gaussianity resembles the effect of a scalar fossil, but is different from that of a vector or a tensor fossil. In particular, lensing does not induce parity-odd BiPoSHs, a unique signature of primordial fields with nonzero spin.

3.2.2 BiPoSH estimators and power spectra

Let us assume a stochastic background of the fossil field, with a particular polarization type α, having a generic power spectrum

$$\left\langle h^{\alpha}_{JM}(K) h^{\alpha'}_{J'M'}(K') \right\rangle = P^{\alpha}_{h}(K) \frac{(2\pi)^3}{K^2} \delta_D(K-K')\, \delta_{JJ'} \delta_{MM'} \delta_{\alpha\alpha'}, \tag{3.19}$$

which is isotropic in the statistical sense. Statistical homogeneity imposes the relations $P^{VE}_h(K) = P^{VB}_h(K) \equiv P^{V}_h(K)$ and $P^{TE}_h(K) = P^{TB}_h(K) \equiv P^{T}_h(K)$. For this reason, we denote $Z = S, V, T$ to take into account scalar, vector and tensor fossils, without differentiating between the two degenerate polarization states for a vector fossil or for a tensor fossil.

Since only one realization of the observable Universe is experimentally accessible, the statistical significance for finding any single BiPoSH coefficient $A^{JM}_{l_1 l_2}$ is expected to be poor; even for a Gaussian random CMB sky, estimators for each $A^{JM}_{l_1 l_2}$ fluctuates about zero with cosmic variance. To improve the significance, we construct the *bipolar auto-/cross-power spectra*

$$C^{J}_{l_1 l_2, l_3 l_4} = \frac{1}{2J+1} \left\langle \sum_{M=-J}^{J} A^{JM}_{l_1 l_2} \left[A^{JM}_{l_3 l_4} \right]^* \right\rangle, \tag{3.20}$$

as we average over all BiPoSH coefficients associated with the same angular scales J, l_1, l_2 but with different orientation M. This is a two-dimensional, spherical harmonic analogy of a four-point correlation ((b) of Figure 3.1), or trispectrum, mediated by the unobservable fossil field. Although statistical isotropy is broken for any given realization of the fossil field, by averaging over all possible realizations of the fossil field, we restore statistical isotropy once again. Therefore, under the full ensemble average those BiPoSH coefficients must be drawn from identical probability distribution. The statistical significance for measuring $C^J_{l_1l_2,l_3l_4}$ above the Gaussian noise is greater than that for measuring any single nonzero BiPoSH coefficient.

A CMB experiment measures the spherical harmonic multipole moments a^T_{lm} as input data. Quadratic estimators may be then constructed to give an unbaised estimate of the BiPoSH departure

$$\widehat{A^{JM}_{l_1l_2}} = \sum_{m_1m_2} (-1)^{m_2} \langle l_1m_1l_2, m_2|JM\rangle \, a^T_{l_1m_1} a^{T*}_{l_2m_2}. \tag{3.21}$$

Under the null hypothesis, the multipole moments a^T_{lm} are independently drawn from Gaussian distributions, so that $\widehat{A^{JM}_{l_1l_2}}$ obeys non-Gaussian statistics. However, with a sufficient number of combinations $a^T_{l_1m_1}$ and $a^T_{l_2m_2}$, the estimators may be Gaussian and independent of each other to good approximation due to the Central Limit Theorem. Then a quadratic estimator may be constructed for the bipolar power spectra

$$\widehat{C^J_{l_1l_2,l_3l_4}} = \frac{1}{2J+1} \sum_{M=-J}^{J} \widehat{A^{JM}_{l_1l_2}} \left[\widehat{A^{JM}_{l_3l_4}}\right]^* - C^{TT}_{l_1} C^{TT}_{l_2} \left(\delta_{l_1l_3}\delta_{l_2l_4} + \delta_{l_1l_4}\delta_{l_2l_3}(-)^{l_1+l_2+J}\right). \tag{3.22}$$

Note that the second term has been subtracted to unbias the estimator from the Gaussian noise expected without fossil effects. A statistically significant detection of non-zero $\widehat{C^J_{l_1 l_2, l_3 l_4}}$ would signal a departure from standard cosmology.

A model-independent way to assess the likelihood of a given model of primordial fossil non-Gaussianity is to parametrize the fossil power spectrum in terms of an overall amplitude $\mathcal{P}_h^Z$ and a fiducial spectrum shape $\tilde{P}_h^Z(K)$,

$$P_h^Z(K) = \mathcal{P}_h^Z \tilde{P}_h^Z(K), \tag{3.23}$$

and similarly parametrize the scalar-scalar-fossil bispectrum in terms of an overall strength $\mathcal{B}_h^Z$ and a fiducial bispectrum shape $\tilde{f}_h^Z(k_1, k_2, K)$ (or a "template"),

$$f_h^Z(k_1, k_2, K) = \mathcal{B}_h^Z \tilde{f}_h^Z(k_1, k_2, K), \tag{3.24}$$

The contribution from a single TAM wave to the BiPoSH coefficients, from Eq. (3.17), is given by

$$A^{JM}_{l_1 l_2}\big|_{h^\alpha_{JM}(K)} = \mathcal{B}_h^Z F^{J,\alpha}_{l_1 l_2}(K) h^\alpha_{JM}(K). \tag{3.25}$$

We have defined the coefficient function

$$\begin{aligned} F^{J,\alpha}_{l_1 l_2}(K) &= -(-i)^J (-1)^{l_1+l_2+P(\alpha)} \frac{16}{\pi} \left(\frac{(2l_1+1)(2l_2+1)}{4\pi} \right)^{\frac{1}{2}} \\ &\times \int k_1^2 dk_1 \, g^T_{l_1}(k_1) \int k_2^2 dk_2 \, g^T_{l_2}(k_2) \, \tilde{f}_h^Z(k_1, k_2, K) \, \mathcal{I}^\alpha_{l_1 l_2 J}(k_1, k_2, K), \end{aligned} \tag{3.26}$$

which can be numerically evaluated with given fiducial bispectrum and concordance cosmology transfer functions. The bipolar power spectra then read

$$C^J_{l_1l_2,l_3l_4} = \frac{\mathcal{A}^Z_h}{(2\pi)^3}\sum_{\alpha\in Z}\int K^2dK\tilde{P}^Z_h(K)F^{J,\alpha}_{l_1l_2}(K)\left[F^{J,\alpha}_{l_3l_4}(K)\right]^* \equiv \mathcal{A}^Z_h\mathcal{F}^{J,Z}_{l_1l_2,l_3l_4}, \tag{3.27}$$

where $\mathcal{A}^Z_h \equiv \mathcal{P}^Z_h(\mathcal{B}^Z_h)^2$ is *the reduced amplitude* of the fossil background. Due to a degeneracy between the amplitude of the fossil-field power spectrum $\mathcal{P}^Z_h$ and the strength of the scalar-scalar-fossil bispectrum $\mathcal{B}^Z_h$, only the reduced amplitude can be directly measured from a bispectrum analysis. Other independent observables, such as higher order correlation functions, are in principle required to break this degeneracy. Of course, a successful inflation model would always make prediction for the two separately, and therefore can be constrained by an analysis of the bispectrum alone. To estimate the reduced amplitude from data, we write

$$\widehat{\mathcal{A}^{J,Z}_{h,l_1l_2,l_3l_4}} = \widehat{C^J_{l_1l_2,l_3l_4}}/\mathcal{F}^{J,Z}_{l_1l_2,l_3l_4}. \tag{3.28}$$

An estimator exists for each pair of possible configurations (J, l_1, l_2) and (J, l_3, l_4), subject to the triangle relation. Without loss of generality, we may restrict to $l_1 \leq l_2$, $l_3 \leq l_4$ and $l_2 \leq l_4$ when $l_1 = l_3$, due to the symmetry of $\widehat{\mathcal{A}^{J,Z}_{h,l_1l_2,l_3l_4}}$ with respect to angular momentum quantum numbers l_i, $i = 1, 2, 3, 4$.

Quadratic estimators $\widehat{\mathcal{A}^{J,Z}_{h,l_1l_2,l_3l_4}}$ are not statistically independent of each other. Even though the off-diagonal elements in the covariance matrix appear sub-dominant compared to the diagonal elements, the effect of off-diagonal elements on likelihood analysis can be

non-negligible.[95] Here, for an order of magnitude estimate of the fossil non-Gaussianity constraints, we shall be content with the assumption that these estimators carry independent information with a large number of pixels from a high-resolution observation, and that they can be weighted with inverse variance to make an optimal estimator with a minimized variance about zero under the null hypothesis,

$$\widehat{\mathcal{A}^{Z}_{h}} = \left\{ \sum_{J} \sum_{(l_1,l_2,l_3,l_4)} \widehat{\mathcal{A}^{J,Z}_{h,l_1l_2,l_3l_4}} \left\langle \left[\widehat{\mathcal{A}^{J,Z}_{h,l_1l_2,l_3l_4}} \right]^2 \right\rangle_0^{-1} \right\} \Big/ \left\{ \sum_{L} \sum_{(l_1,l_2,l_3,l_4)} \left\langle \left[\widehat{\mathcal{A}^{J,Z}_{h,l_1l_2,l_3l_4}} \right]^2 \right\rangle_0^{-1} \right\}. \tag{3.29}$$

Here $\sum_{(l_1,l_2,l_3,l_4)}$ sums over all possible combinations of multipole quantum numbers. We shall assume that the distribution of this optimal estimator is approximately Gaussian, so that its variance, evaluated to be

$$(\sigma^{Z}_{\mathcal{A}})^{-2} \equiv \left\langle \left[\widehat{\mathcal{A}^{Z}_{h}} \right]^2 \right\rangle_0^{-1} = \frac{1}{8} \sum_{J} \sum_{l_1l_2l_3l_4} \frac{2J+1}{C^{TT}_{l_1} C^{TT}_{l_2} C^{TT}_{l_3} C^{TT}_{l_4}} \mathcal{F}^{J,Z}_{l_1l_2,l_3l_4} \left[\mathcal{F}^{J,Z}_{l_1l_2,l_3l_4} \right]^*, \tag{3.30}$$

provides a good characterization of the scatter due to cosmic variance. This roughly gives the cosmic variance limitation for detecting a primordial scalar-scalar-fossil bispectrum with statistical significance. In practice, instrumental noise and foreground contamination would further degrade the estimated sensitivity. Assessment of these challenges go beyond the scope of our discussion.

3.2.3 Numerical study

As we have pointed out, a primordial scalar-scalar-fossil bispectrum dominated in the squeezed configuration is capable of producing modulation effects in the CMB on very large angular scales, and therefore can be distinguished from gravitational lensing. We now consider such a squeezed type bispectrum by assuming

$$\tilde{f}_h^Z(k_1, k_2, K) = (k_1 k_2)^{-3/2}. \tag{3.31}$$

Note that in the squeezed limit $k_1 \approx k_2 \gg K$ although we prefer a symmetric form. We furthermore consider fossil power spectrum in the scale-invariant limit,

$$\tilde{P}_h^Z(K) = 1/K^{-3}, \tag{3.32}$$

if no special physical scale other than the Hubble scale plays a role during the slow-roll phase.

Gravitational waves amplified during the slow-roll phase provide an explicit example of tensor fossil for which the two fiducial forms being assumed here are physically realized. Note that for gravitational waves, the bispectrum strength $\mathcal{B}_h^Z$ is expected to be order unity multiplied by the potential power spectrum $k^3 P_\Phi(k) \sim 10^{-8}$, while the power spectrum amplitude $\mathcal{P}_h^Z$ is currently constrained to be $\lesssim 10^{-10}$ by measurement of the B-mode polarization patterns in the CMB; therefore, the gravitational wave signature $A_h^Z \lesssim 10^{-26}$, for standard inflationary scenario, is unfortunately way below current or future CMB detection

limit. Possible detection of this tiny signature will have to wait for future measurement of 21cm fluctuations in the Dark Ages.

We assume the concordance ΛCDM cosmology with cosmological parameters from the *WMAP*+BAO+H_0 best fit.[96] Since our analysis was done, values for the cosmological parameters have changed slightly according to latest data. However, those updates are unimportant to our order-of-magnitude forecasts. We further approximate by taking a perfectly scale-invariant power spectrum $n_s = 1$ for the primordial potential perturbations. Using the publicly available code `CAMB`,[97] we consider a high resolution experiment with $l_{\rm max} = 3000$ and tabulate radiation transfer functions needed for Eq. (3.26).

For inflation, it is natural to extrapolate the fossil power spectrum to very long wavelength or very small wave number K. For a scale invariant fossil power spectrum, a logarithmic singularity is present, a problem to be cured by an infrared cut-off. We first comment that the infrared cut-off must in general arise physically, as a fossil perturbation with a wavelength far longer than the size of the visible Universe is indistinguishable from a spatially constant configuration, which should induce no physical effect.[98,99] In our case here, in the TAM formalism the infrared end is geometrically cut off for parity-even modes with $J > 2$ and all parity-odd modes, since the TAM wave function essentially vanishes within the observable Universe, if the wavelength $\sim 1/K$ extends beyond the Hubble scale. The only exception is the parity-even $J = 2$ mode, whose wavefunction is finite at the spatial origin $\mathbf{x} = 0$. Since the logarithmic divergence is mild, we circumvent the problem by simply truncating the fossil wave number at $K_{\rm min} = 5 \times 10^{-6}{\rm Mpc}^{-1}$, corresponding to a reasonable length scale slightly larger than the size of the present horizon.

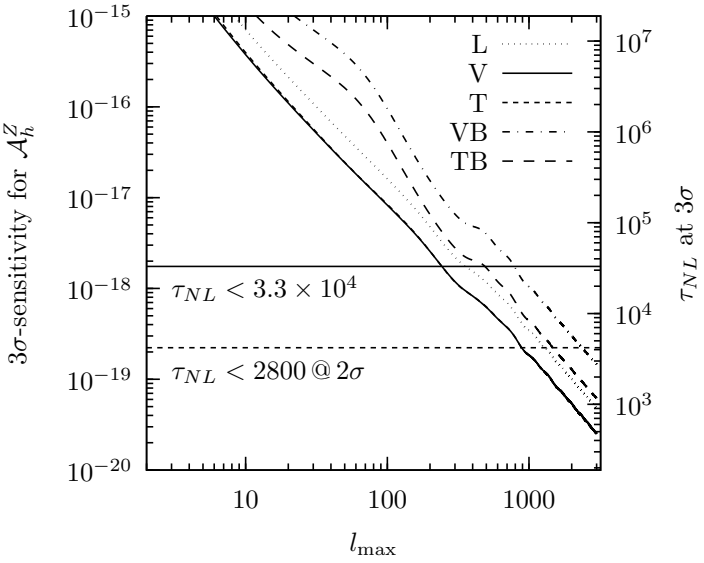


Figure 3.5: Three-σ sensitivity for the reduced amplitude $\mathcal{A}_h^Z$ versus the angular resolution $l_{\max}$ from Ref.[62] Only BiPoSHs with $J \leqslant 5$ are included. Our results can be related to the primordial trispectrum parameter τ_{NL} in the local model of non-Gaussianity through $\mathcal{A}_h^L = 5.3 \times 10^{-23} \tau_{NL}$.[100] Also shown on the plot are the *WMAP* 5-year constraint (horizontal solid), and the constraint from *Planck 2013* results (horizontal dashed). The analysis of the *Planck 2015* results on primordial trispectrum uses a somewhat different parameterization than the τ_{NL} parameter.[83] Since the constraining power is not anticipated to be significantly improved from the 2013 results to the 2015 results, we are content with a comparison with the 2013 results.

We found that large-scale modulation $J \leq 5$ already dominates the signal-to-noise of the fossil BiPoSHs. On the other hand, if we let $l_{\max}$ to be the smallest angular scale that is dominated by cosmological signal, the signal-to-noise scales quickly with improvement in the resolution $\sim l_{\max}^2$, in proportion with the total number of pixels accessible by the experiment. This scaling law can be easily verified in the Sachs-Wolfe limit above the sound horizon scale: The radiation transfer function becomes purely geometrical $g_l^T(k) = -j_l(kr^\star)/3$ with $r^\star \approx$ 14 Gpc being the comoving distance to the last scattering surface in concordance cosmology. For fixed modulation scale J and fluctuation scales l_1, l_2 (in the squeezed configuration $J \ll l_1 \approx l_2 \approx l_{12}$ and $J \ll l_3 \approx l_4 \approx l_{34}$), the signal squared decreases as $l_{12}^{-3} l_{34}^{-3}$,

while the noise squared decreases more rapidly as $l_{12}^{-4}l_{34}^{-4}$, so that each estimator $\widehat{\mathcal{A}^{J,Z}_{h,l_1l_2,l_3l_4}}$ contributes a signal-to-noise squared $\propto l_{12}l_{34}$. The number of triangle configurations subject to $|l_1 - l_2| < J$ and $|l_3 - l_4| < J$ is proportional to $J\sum_{l_{12}\leqslant l_{\rm max}}\sum_{l_{34}\leqslant l_{\rm max}}$, leading to a cumulative signal-to-noise squared $\mathcal{S}^2 \propto J\sum_{l_{12}\leqslant l_{\rm max}}\sum_{l_{34}\leqslant l_{\rm max}} l_{12}l_{34} \propto l_{\rm max}^4$.

We plot the 3σ detection limit for the reduced fossil amplitude in Figure 3.5 as a function of the angular resolution $l_{\rm max}$, which confirms the $l_{\rm max}^{-2}$ scaling behavior. The case of scalar, vector, and tensor fossil field is plotted separately. When both parity-even and parity-odd BiPoSHs for a vector fossil or a tensor fossil are included, the detection sensitivity is roughly twice as good as the case of a scalar fossil, due to doubling of the polarization states. We also plot the sensitivities including only parity-odd BiPoSHs (with a compromise in overall sensitivity though), which would not be generated via lensing by foreground gravitational potential. Those are the smoking gun of vector or tensor new physics in the primordial Universe. We find that our estimators are more sensitive to parity-odd BiPoSHs generated by a tensor fossil field than that generated by a vector fossil field.

Our results show that the *Planck* experiment with $l_{\rm max} \sim 2500$ is capable of probing reduced fossil amplitude down to $A_h^Z \approx 5\times10^{-20} - 10^{-19}$, comparable to redshift LSS surveys with a large comoving volume $V \sim 60$ Gpc$^3h^{-3}$ but a moderate resoltuion $k_{\rm max} = 0.1h$Mpc^{-1}.

3.2.4 Relation to primordial trispectrum τ_{NL}

For the case of a scalar fossil field, the effective trispectrum ((b) of Figure 3.1) mediated by an internal fossil field mode is conventionally parametrized terms of the τ_{NL} parameter, which is defined through the *connected* part of the primordial potential trispectrum

$$\langle \Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\Phi(\mathbf{k}_4)\rangle = (2\pi)^3 \delta_D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3+\mathbf{k}_4) T_\Phi(k_1,k_2,k_3,k_4), \tag{3.33}$$

where

$$T_\Phi(k_1,k_2,k_3,k_4) = \frac{1}{2}\frac{25}{9}\tau_{NL}\left\{P_\Phi(k_1)P_\Phi(k_2)P_\Phi(|\mathbf{k}_1+\mathbf{k}_3|) + 23\,\text{cyclic}\right\}. \tag{3.34}$$

If we put together two scalar-scalar-fossil triangles of squeezed configuration, which share the same third side of fossil mode, a connected potential trispectrum is formed, which is given by

$$\begin{aligned} T_\Phi(k_1,k_2,k_3,k_4) &= f_h^\alpha(k_1,k_2,K) f_h^\alpha(k_3,k_4,K) \sum_\alpha \epsilon_{ij}^\alpha(\mathbf{K}) \hat{k}_1^i \hat{k}_2^j \epsilon_{mn}^{*\alpha}(\mathbf{K}) \hat{k}_3^m \hat{k}_4^n P_h^\alpha(|\mathbf{k}_1+\mathbf{k}_2|) \\ &\approx \left(\frac{f_h^\alpha}{P_\Phi}\right)^2 \left(\frac{A_h^\alpha}{P_\Phi}\right) P_\Phi(k_1)P_\Phi(k_2)P_\Phi(|\mathbf{k}_1+\mathbf{k}_2|) \left[\sum_\alpha \epsilon_{ij}^\alpha(\mathbf{K}) \hat{k}_1^i \hat{k}_2^j \epsilon_{mn}^{*\alpha}(\mathbf{K}) \hat{k}_3^m \hat{k}_4^n\right] \\ &= \frac{1}{2}\frac{2\mathcal{A}_h^\alpha}{P_\Phi^3} P_\Phi(k_1)P_\Phi(k_2)P_\Phi(|\mathbf{k}_1+\mathbf{k}_2|) \left[\sum_\alpha \epsilon_{ij}^\alpha(\mathbf{K}) \hat{k}_1^i \hat{k}_2^j \epsilon_{mn}^{*\alpha}(\mathbf{K}) \hat{k}_3^m \hat{k}_4^n\right]. \end{aligned} \tag{3.35}$$

Since the polarization sum in the square brackets evaluates to unity for the case of scalar fossil, we derive for the case of a scalar fossil field the conversion between our reduced fossil

amplitude and τ_{NL},

$$\mathcal{A}_h^L = 5.3 \times 10^{-23} \tau_{NL}, \tag{3.36}$$

for the case of a scale-invariant primordial potential power spectrum $P_\Phi(k) \approx 1.72 \times 10^{-8}\, k^{-3}$. Through this conversion, our results for scalar fossil field is consistent with previous forecasts on local-type primordial trispectrum.[100,101]

3.3 Summary

We point out that existing analysis on primordial trispectrum τ_{NL} have not constrained fossil fields of vector or tensor type. This is because the estimator for local-type τ_{NL}, Eq. (3.33) and Eq. (3.34), picks up a collapsed four-point function that is independent of the relative azimuthal angle $\Delta\varphi$ between the two squeezed triangular configurations, as shown in Figure 3.6, while a vector or a tensor fossil field is expected to induce higher order harmonic dependence on $\Delta\varphi$. This opens up the possibility of searching for non-trivial primordial potential trispectrum mediated by new vector or tensor fields with current high-resolution CMB experiments. The search will be complimentary to conventional searches for a primordial scalar-scalar-scalar bispectrum, which only involves the adiabatic mode during inflation and needs not to be mediated by a fossil field. Possibly due to discrete symmetry or supersymmetry, it is conceivable that there might be a large primordial scalar trispectrum in the absence of a significant scalar bispectrum.[102]

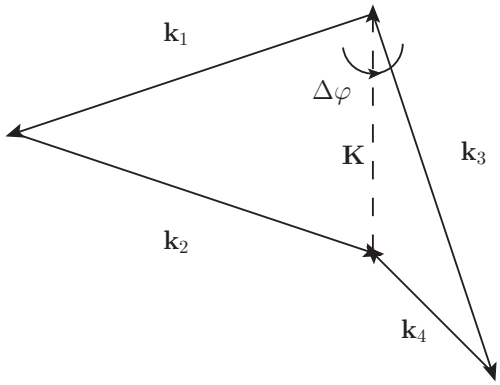


Figure 3.6: How a primordial potential (solid) trispectrum depends on the azimuthal separation $\Delta\varphi$ between the two triangle configurations is determined by the spin of the intermediate fossil mode (dashed).

Searching for possible trispectrum arising from "exchanged diagrams" due to a scalar-scalar-new-field coupling has so far been left out for future study even in the most recent data analysis of the *Planck* experiment.[83] If done, this will put constraints on multi-field inflationary scenarios.[70,71,103–105]

Since the primordial potential perturbation induces both temperature fluctuations and polarization in the CMB, a primordial scalar-scalar-fossil bispectrum is expected to break statistical isotropy not only in the temperature map, but in the polarization map as well. Just like the case of temperature, fossil fields on large angular scales might modulate the two-point statistics of linear polarization Stokes parameters on smaller scales. Sourced by scalar perturbations, this applies to parity-even E-type polarization anisotropies but not for parity-odd B-type polarization anisotropies. Our analysis on temperature anisotropy can be directly generalized to polarization anisotropy using a generalized bipolar spherical harmonic formalism to account for spin-two Stokes parameters.[106,107] Including polarization anisotropies adds to the total number of independent harmonic modes and hence decreases

the cosmic variance limitation on non-Gaussianity measurement. However, the improvement on the overall signal-to-noise is expected to be at most a factor of few, depending on the polarization map resolution and the specifics of foreground or instrumental noise, and therefore will only be important for further investigation should a suspicious primordial signature surface.

Chapter 4

Gravitational Waves and CMB Lensing

Resulting from a variety of late-time effects, the observed fluctuations in the CMB do not faithfully represent the initial perturbations as generated by the inflation mechanism. An important category of late-time effects arise from general relativistic kinetic corrections to photon propagation in a perturbed spacetime. After all, in cosmology and astronomy we rely heavily on the reception of electromagnetic quanta to make inference about the physical property of distant sources. Therefore, in order to be able to discover any new physics in the initial condition, such as what has been discussed in Chapter 3, thorough understanding and modelling of the late-time effects in the CMB are imperative. These effects in turn provide an avenue to look for new phenomena in the late-time Universe, such as the existence of primordial gravitational waves.

A photon carries three pieces of information as it travels through spacetime. First, it has a fixed (angular) frequency ω, corresponding to a fixed energy $E = \hbar\omega$. Second, it propagates along a fixed direction $\hat{\mathbf{n}}$, which determines the momentum the photon carries $\mathbf{p} = \hbar\omega\hat{\mathbf{n}}$. Finally, the photon is either in one of the two independent linear polarization states, both perpendicular to the direction of propagation, or in a superposition.

Note that according to special relativity such measurement of the photon state is always reference-frame dependent. The outcome of the measurement depends on the motion of both the emitting source and the observer. For instance, relative velocity between the source and the observer is known to induce a relativistic Doppler shift in the photon frequency, or equivalently, its energy. Moreover, change to the measured direction of propagation due to such relative motion is referred to as the relativistic aberration.[108] In addition, the orientation of photon polarization may change under aberration. In a fully general relativistic treatment, photon energy undergoes further redshift or blueshift when it climbs out of local gravitational potential. This Sachs-Wolfe effect is responsible for the large-scale temperature anisotropies we observe on the surface of last scatter.[109] While local motion and local gravitational potential are indeed important effects to account for in CMB physics, that will not be the focus of study in this Chapter. Instead, we consider non-local line-of-sight integral effects on photon propagation as a result of spacetime lumpiness.

Achromatic redshift/blueshift of photon energy occurs if the photon traverses a region with time-dependent gravitational potential. This is called the Integrated Sachs-Wolfe (ISW) effect[109] or Rees-Sciama effect[110] in a more general sense, a major source of CMB temperature anisotropies on large angular scales resulting from the decaying large-scale

gravitational potential during the recent epoch of Dark Energy domination.[111] More generally, any perturbation to spacetime metric may generate ISW effect. For instance, gravitational waves distort spatial metric and may account for a fraction of the large-scale temperature anisotropies in the CMB if they were produced in the early Universe.

Integrated deflection of photon direction of propagation is usually referred to as gravitational lensing, a famous prediction of general relativity. The concept has been proposed long ago in the study of gravitational deflection of background star-light near massive celestial objects. On cosmological scales, smooth but inhomogeneous mass distribution can also deflect the paths of CMB photon, in addition to compact massive objects such as galaxies and clusters.[112,113] Since the image distortion in this regime is subtle, this is usually called weak gravitational lensing, in contrast to strong gravitational lensing which typically produces multiple images. This phenomenon in the CMB has a wide scope of applications on extracting cosmological informations.[89,90,114,115] Besides, vector or tensor metric perturbations, possibly induced by cosmic strings,[116] or originating from inflation, may lens the CMB as well.[86,117–119]

The statistics of polarization anisotropies in the CMB is also affected by gravitational lensing by intervening large scale structure.[90,120] This is because inhomogeneous gravitational deflection geometrically distort the observed polarization pattern on the sky. However, apart from deflection, another aspect of the physics had been largely neglected in the literature until our recent work,[58] which is the fact that the plane of linear polarization rotates about the direction of propagation as it is being parallel transported through a curved spacetime. This aspect is not important for weak lensing by scalar gravitational potential,

because this polarization rotation is zero for a linearized calculation under the *Born approximation*, which is usually sufficient for weakly perturbed spacetime metric. However, it has to be consistently accounted for when calculating weak lensing by vector or tensor metric perturbations, e.g. when searching for evidence of vorticities or gravitational waves in the early Universe.

A new contribution to the study of weak lensing physics is the elucidation of this polarizaton rotation effect, based on the published work of Ref.[58] Although this will be the highlight of this Chapter, for a self-contained presentation we will first lay down the complete harmonic-space formalism of weak lensing by general metric perturbations, for both temperature and polarization observables. We first discuss in Section 4.1 the transportation of photon energy-momentum and polarization through a weakly perturbed spacetime using general relativity. In Section 4.2, we then formulate the general effect of weak lensing on harmonic-space observables using the TAM formalism. Finally, Section 4.3 is devoted to the case study of weak lensing by gravitational waves, highlighting on the polarization rotation effect. We close this Chapter by a brief summary Section 4.4 in the end.

4.1 Transportation in perturbed spacetime

In cosmological applications, typically the photon wavelength is much shorter than the length scale associated with spacetime curvature. In this geometrical optics limit, the propagation of photon is described by geodesic transportation of the photon covariant 4-momentum vector and polarization vector. This shall lay down the basis of our discussion.

For simplicity, we will first be content with a calculation for a perturbed Minkowskian spacetime, rather than for a perturbed universe of homogeneous expansion. Although in many astrophysical applications, weak lensing in a perturbed Minkowskian spacetime is sufficient, there are many cosmological contexts in which distances comparable to the horizon scale (for instance the distance between the last scattering surface and us) are involved. For those cases, the calculation needs to be carried out in a perturbed FLRW spacetime. Nevertheless, we will see that eventually the result for a perturbed Minkowskian spacetime can be straightforwardly generalized to the desired result for a perturbed FLRW spacetime, due to invariance of photon geodesic motion under a conformal transformation, which properly accounts for the effect of the cosmological scale factor.

We will also be content with a calculation valid to linear order in metric perturbations. This approximation is practically valid in all cosmological contexts. Let us consider the following perturbed spacetime

$$ds^2 = -\left[1+2a(\mathbf{x},t)\right]dt^2 + 2b_i(\mathbf{x},t)dtdx^i + \left[\delta_{ij}+h_{ij}(\mathbf{x},t)\right]dx^i dx^j. \tag{4.1}$$

Perturbation has been parametrized in a general gauge by a scalar $a(\mathbf{x},t)$, a vector $b_i(\mathbf{x},t)$, and a rank-two tensor $h_{ij}(\mathbf{x},t)$. Useful algebraic results are collected in Appendix A.

We define stationary observers at each spatial location, who are stationary with respect to the *unperturbed* metric. They are different from free-falling observers; those have peculiar velocities with respect to the background metric as they fall into gravitational potentials.

We make theoretical prediction for what is measured by those stationary observers. We

do not choose a free-falling observer, because the effect of the observer's peculiar velocity is often treated separately. Stationary observers define a time direction

$$e^{\mu}_{(0)} = \{1/\sqrt{-g_{00}}, \vec{0}\} = \{1 - a, \vec{0}\}, \tag{4.2}$$

They then define three orthogonal spatial vectors

$$e^{\mu}_{(i)} = \left\{ b_i, \delta^j{}_i - \frac{1}{2} h^j{}_i \right\}. \tag{4.3}$$

This choice of $e^{\mu}_{(i)}$ are *non-rotating*. They differ minimally from the natural tetrad vectors in an unperturbed spacetime as the deviation is linear in metric perturbations. We will see that this point becomes crucial when we discuss polarization rotation.

We will use a dot to denote a derivative with respect to time t, and use ∂_i for a derivative with respect to the spatial coordinate $\mathbf{x}$.

4.1.1 Energy and momentum

Consider a photon with 4-momentum p_μ, as shown in Figure 4.1. It has polarization tensor $P_{\mu\nu}$ normalized to $\delta^{\mu\nu} P_{\mu\nu} = 1$. Upon observation, suppose the photon has observed energy ω_o and observed direction $\hat{n}^i$ in the sky. The photon 4-momentum then reads

$$p^{\mu}(t = t_o) = \omega_o e^{\mu}_{(0)} - \omega_o \hat{n}^i e^{\mu}_{(i)}, \tag{4.4}$$

so that

$$\frac{1}{\omega_o}p^0(t=t_o) = \frac{1}{\omega_o}p^0_o = 1 - a_o - \hat{n}^i b_{o,i}, \tag{4.5}$$

$$\frac{1}{\omega_o}p^i(t=t_o) = \frac{1}{\omega_o}p^i_o = -\hat{n}^i + \frac{1}{2}\hat{n}^j (h_o)^i{}_j. \tag{4.6}$$

Of course, if there are no metric perturbations, the zeroth order photon 4-momentum is given by $p^{0(0)} = \omega_o$ and $p^{i(0)} = -\omega_o \hat{n}^i$ ("$_o$" stands for observing time and location, and "$_e$" stands for emission time and location).

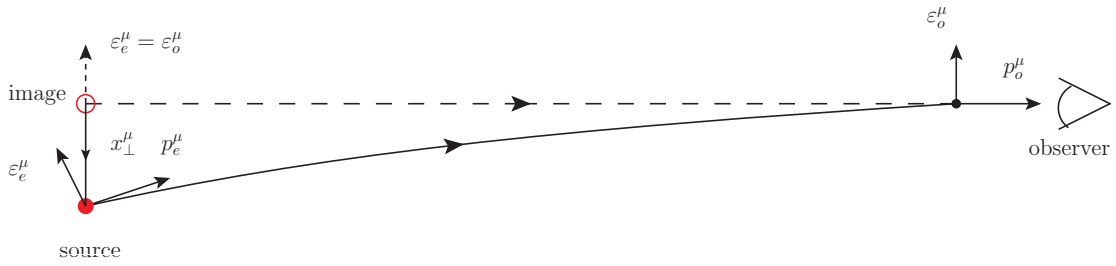


Figure 4.1: Gravitational lensing of light from a distant source. The physical trajectory is shown with the solid line. The undeflected trajectory, which would be true in an unperturbed spacetime, is shown with the dashed line. It is used to infer the image position.

The photon 4-momentum is geodesic transported along the null trajectory,

$$\frac{dp^\mu}{d\lambda} = -\Gamma^\mu_{\alpha\beta}\, p^\alpha p^\beta, \tag{4.7}$$

parametrized by the affine parameter λ. We have $p^0 = dt/d\lambda$. Since $dp^\mu/d\lambda$ is at least first order, we only need to substitute in $dt/d\lambda = \omega_o$, so that

$$\omega_o\, \dot{p}^\mu = -\Gamma^\mu_{\alpha\beta}\, p^\alpha p^\beta. \tag{4.8}$$

Neglecting terms beyond linear order in perturbations, we then have equations for the first-order correction due to metric perturbations,

$$\omega_o \, \dot{p}^{\mu(1)} \;=\; -\Gamma^{\mu}_{\alpha\beta} \, p^{\alpha(0)} p^{\beta(0)}. \tag{4.9}$$

A decomposition into the temporal and the spatial component yields

$$\frac{1}{\omega_o}\dot{p}^{0(1)} \;=\; -\Gamma^{0}_{00} + 2\Gamma^{0}_{0i}\hat{n}^i - \Gamma^{0}_{ij}\hat{n}^i\hat{n}^j, \tag{4.10}$$

$$\frac{1}{\omega_o}\dot{p}^{i(1)} \;=\; -\Gamma^{i}_{00} + 2\Gamma^{i}_{0j}\hat{n}^j - \Gamma^{i}_{jk}\hat{n}^j\hat{n}^k. \tag{4.11}$$

In the Born approximation, one integrates along the *unperturbed* line of sight $dt = -dr$ (which is sufficient at linear order), where r parametrizes the distance along the line of sight. Note, however, that an appropriate boundary condition must be specified at the observer's location.

(Integrated) Sachs-Wolfe effect: Consider an inertial observer at rest relative to the background (Minkowskian) metric. She measures the photon energy to be

$$\omega \;=\; -g_{\mu\nu} \, p^{\mu} \, e^{\nu}_{(0)} = \omega_o \left(1 + a + \frac{p^{0(1)}}{\omega_o} + \hat{n}^i b_i\right), \tag{4.12}$$

The requirement that the observer measures an energy ω_o fixes the boundary condition $p^{0(1)} = -\omega_o \left(a_o + \hat{n}^i b_{o,i}\right)$ at the observer's location. The *fractional* perturbation in photon energy (integrated Sachs-Wolfe effect) is then obtained by evaluating Eq. (4.12) at the

source location,

$$\begin{aligned}\frac{\omega_e}{\omega_o}-1 &= a_e+\hat{n}^i b_{e,i}+\frac{p^{0(1)}}{\omega_o} \\ &= a_e-a_o+n^j\left(b_{e,i}-b_{o,i}\right)-\int_o^e dr\left[-\dot{a}+2\hat{n}^i\partial_i a+\hat{n}^i\hat{n}^j\partial_i b_j-\frac{1}{2}\hat{n}^i\hat{n}^j\dot{h}_{ij}\right].\end{aligned} \tag{4.13}$$

Here and thereafter, integrands of line-of-sight integrals are understood to be evaluated on the *past light cone* $d/dr = -\partial_t + \hat{n}^i\partial_i$. A familiar case is lensing by scalar gravitational potentials, for which we choose the Newtonian gauge $a = \Phi$, $b_i = 0$ and $h_{ij} = -2\Psi\delta_{ij}$. Then the redshift is

$$\begin{aligned}\frac{\omega_e}{\omega_o}-1 &= -\left(\Phi_o-\Phi_e\right)-\int_o^e dr\left[-\dot{\Phi}+2\hat{n}^i\partial_i\Phi+\dot{\Psi}\right] \\ &= -\left(\Phi_e-\Phi_o\right)-\int_o^e dr\left(\dot{\Phi}+\dot{\Psi}\right).\end{aligned} \tag{4.14}$$

The first term gives the Sachs-Wolfe effect, and the second term gives the integrated Sachs-Wolfe effect whenever the gravitational potentials are non-static.

Deflection: The solution for 3-momentum is given by

$$\begin{aligned}\frac{p^i}{\omega_o} &= -\hat{n}^i+\frac{1}{2}\hat{n}^j(h_o)^i{}_j \\ &\quad -\int_o^e dr\left[-\partial^i a-\dot{b}^i+\hat{n}^j\left(\partial_j b^i-\partial^i b_j\right)+\hat{n}^j\dot{h}^i{}_j-\frac{1}{2}\hat{n}^j\hat{n}^k\left(\partial_j h^i{}_k+\partial_k h^i{}_j-\partial^i h_{jk}\right)\right].\end{aligned} \tag{4.15}$$

To find the source position, we use

$$\frac{dx^i}{d\lambda}=p^0\frac{dx^i}{dt}=p^i, \tag{4.16}$$

and then convert between the affine parameter λ and coordinate time t. The part transverse to $\hat{n}^i$ can be obtained by contracting with the projection tensor $\Pi^i{}_j = \delta^i_j - \hat{n}^i\hat{n}_j$, and we obtain

$$\begin{aligned}\frac{dx^i_\perp}{dt} &= \frac{1}{2}\Pi^i{}_j\hat{n}^k(h_o)^j{}_k \\ &\quad -\Pi^i{}_j\int_o^e dr\left[-\partial^j a - \dot{b}^j + \hat{n}^k\left(\partial_k b^j - \partial^j b_k\right) + \hat{n}^k\dot{h}^j{}_k - \frac{1}{2}\hat{n}^k\hat{n}^l\left(\partial_k h^j{}_l + \partial_l h^j{}_k - \partial^j h_{kl}\right)\right].\end{aligned} \tag{4.17}$$

Since the transverse displacement is at least first order in metric perturbation, the source light-cone effect (namely the observed time of the source is perturbed) can be neglected, which justifies a direct integration over the unperturbed line of sight in order to find $x^i_\perp$. After rather tedious algebra, we find the transverse angular displacement $\theta^i = x^i_\perp/r_e$ to be given by

$$\begin{aligned}\theta^i &= \Pi^i{}_j\left\{\left(-b^j_o + \frac{1}{2}\hat{n}^k(h_o)^j{}_k\right)\right. \\ &\quad \left.+\int_o^e dr\left[\left(1-\frac{z}{z_e}\right)\left(-\partial^j a - \hat{n}^k\partial_j b_k + \frac{1}{2}\hat{n}^k\hat{n}^l\partial^j h_{kl}\right) + \frac{1}{z_e}\left(b^j - \hat{n}^k h^j{}_k\right)\right]\right\},\end{aligned}$$

in agreement with e.g. Ref.[121] Inserting the special case of gravitational potential $a = \Phi$, $b_i = 0$ and $h_{ij} = -2\Psi\delta_{ij}$, we reproduce the familiar result for weak lensing deflection

$$\theta^i = -2\int_o^e dr\left(1-\frac{r}{r_e}\right)\Pi^i{}_j\,\partial^j\Psi. \tag{4.18}$$

4.1.2 Polarization

The photon polarization state can be specified by a polarization vector ε^μ [1], which is parallel transported along the photon's geodesic $p^\nu \nabla_\nu \varepsilon^\mu = 0$. The on-shell condition imposes $\varepsilon^\mu p_\mu = 0$. This is forced along the geodesic because poth ε^μ and p^μ are parallel-transported along it.

The polarization vector ε^μ still has one gauge degree of freedom,

$$\varepsilon^\mu \longrightarrow \varepsilon^\mu + \alpha\, p^\mu, \tag{4.19}$$

where α is any coefficient. This preserves the on-shell condition because the photon is massless $p^\mu p_\mu = 0$. This introduces an ambiguity when interpreting the result for ε^μ.

This gauge redundancy can be eliminated by a screen-projection procedure.[122] Consider an inertial observer at any *single* space-time point, the observer defines a time direction $e^\mu_{(0)}$. The 4-momentum p^μ projected onto the perpendicular spatial hyper-plane then points toward a spatial direction (direction of propagation),

$$\hat{n}^\mu \;=\; \frac{p^\mu}{\omega} - e^\mu_{(0)}, \tag{4.20}$$

where ω is the photon energy measured by the observer. The *screen projection tensor*,

[1] Strictly speaking, the most general polarization state is specified by a two-by-two polarization tensor, although when the photon is in a pure state, the polarization tensor is the direct product of a doubling of the polarization vector associated with that pure state. In that case, flipping the sign of the polarization vector has no physical consequences.

defined as

$$S^{\mu}{}_{\nu} = g^{\mu}_{\nu} + e^{\mu}_{(0)} e_{(0),\nu} - \hat{n}^{\mu}\hat{n}_{\nu}, \tag{4.21}$$

projects any vector onto a two-dimensional subspace perpendicular both to the time direction and the direction of propagation. When acting on ε^{μ}

$$S^{\mu}{}_{\nu}\,\varepsilon^{\nu} = \varepsilon^{\mu} + \frac{p^{\mu}}{\omega} e^{\nu}_{(0)}\,\varepsilon_{\nu}, \tag{4.22}$$

it exactly eliminates the gauge piece in Eq. (4.19). This procedure can be in particular carried out at the observer's location and at the source location.

We are now in a position to solve ε^{μ} along the photon trajectory. The initial condition can be written in terms of the observer's tetrads,

$$\varepsilon^{\mu}(t=0) = \varepsilon^{(a)}(t=0) e^{\mu}_{(a)} \equiv \varepsilon^{(a)}_{o} e^{\mu}_{(a)}, \qquad a = 0,1,2,3. \tag{4.23}$$

Using Eq. (4.2) and Eq. (4.3), the expressions for $\varepsilon^{\mu}(t=0)$ read

$$\varepsilon^{0}(t=0) = \varepsilon^{0}_{o} = \varepsilon^{(0)}_{o} - \varepsilon^{(0)}_{o} a_{o} + \varepsilon^{(i)}_{o} b_{o,i}, \tag{4.24}$$

$$\varepsilon^{i}(t=0) = \varepsilon^{i}_{o} = \varepsilon^{(i)}_{o} - \frac{1}{2}\varepsilon^{(j)}_{o} (h_{o})^{i}{}_{j}. \tag{4.25}$$

The zeroth-order solution (solution in Minkowski spacetime) is simply

$$\varepsilon^{0(0)} = \varepsilon^{(0)}_{o}, \qquad \varepsilon^{i(0)} = \varepsilon^{(i)}_{o}. \tag{4.26}$$

The geodesic equation for polarization reads

$$\frac{d\varepsilon^{\mu}}{d\lambda} = p^0 \frac{d\varepsilon^{\mu}}{dt} = -\Gamma^{\mu}_{\alpha\beta}\, p^{\alpha}\varepsilon^{\beta}. \tag{4.27}$$

The first-order part of the equation is

$$\frac{d\varepsilon^{0(1)}}{dt} = -\Gamma^0_{00}\,\varepsilon^{(0)}_o - \Gamma^0_{0i}\,\varepsilon^{(i)}_o + \Gamma^0_{i0}\,\hat{n}^i\varepsilon^{(0)}_o + \Gamma^0_{ij}\,\hat{n}^i\varepsilon^{(j)}_o, \tag{4.28}$$

$$\frac{d\varepsilon^{i(1)}}{dt} = -\Gamma^i_{00}\,\varepsilon^{(0)}_o - \Gamma^i_{0j}\,\varepsilon^{(j)}_o + \Gamma^i_{j0}\,\hat{n}^j\varepsilon^{(0)}_o + \Gamma^i_{jk}\,\hat{n}^j\varepsilon^{(k)}_o. \tag{4.29}$$

A straightforward integrations along the *unperturbed* line of sight gives

$$\begin{aligned}
\varepsilon^0_e &= \varepsilon^{(0)}_o - \varepsilon^{(0)}_o a_e + \frac{1}{2}\varepsilon^{(i)}_o\,(b_{o,i}+b_{e,i}) + \varepsilon^{(i)}_o \int_o^e dr\left[\partial_i a + \frac{1}{2}\dot{b}_i + \frac{1}{2}\hat{n}^j\partial_i b_j - \frac{1}{2}n^j\dot{h}_{ij}\right],\\
\varepsilon^i_e &= \varepsilon^{(i)}_o - \frac{1}{2}\varepsilon^{(0)}_o\left(b^i_e - b^i_o\right) - \frac{1}{2}\varepsilon^{(j)}_o (h_e)^i{}_j + \varepsilon^{(0)}_o\int_o^e dr\left[\partial^i a + \frac{1}{2}\dot{b}^i + \frac{1}{2}\hat{n}^j\partial^i b_j - \frac{1}{2}\hat{n}^j\dot{h}^i{}_j\right]\\
&\quad + \frac{1}{2}\varepsilon^{(j)}_o\int_o^e dr\left[\partial_j b^i - \partial^i b_j - \hat{n}^k\left(\partial_j h^i{}_k - \partial^i h_{jk}\right)\right].
\end{aligned} \tag{4.30}$$

Polarization rotation: What do we really mean by "the polarization plane rotates in a perturbed spacetime"? How would we know whether the polarization plane has rotated with respect to some reference direction at all?

In flat spacetime it is unambigious to compare the direction of a 4-vector at two separate locations — the vector can be decomposed in terms of a set of tetrad vectors — its " spatial orientation" is given by the expansion coefficients, $\varepsilon^{(a)}$ in Eq. (4.23), for the three spatial tetrad vectors $a = 1, 2, 3$. In a flat spacetime, all points naturally share the same set of tetrad vectors that can be trivially translated back and forth, so any rotation is signified

by a change in the expansion coefficients. The flat spacetime case would correspond to the unperturbed trajectory in Figure 4.1, along which both ε^μ and tetrad vectors are transported in the same way (in particular, three unit vectors in the x-, y- and z-directions at one point would correspond to the three unit vectors at another point respectively). Therefore, there is *no* polarization rotation when photon is propagating in flat spacetime.

The situation changes in a perturbed spacetime, in which a photon travels along a deflected trajectory in Figure 4.1. At both ends of the trajectory, ε^μ may be expanded in terms of the local set of tetrad vectors. However, the tetrad vectors differ at the observer's location and the source location, the "minimal" choice given by Eqs. (4.2)–(4.3), are not related by geodesic transport, and therefore a non-zero change in the corresponding expansion coefficients $\varepsilon^{(i)}$ signals a physical rotation of the polarization plane due to curved spacetime geometry. If they otherwise are, geodesic transportation of the polarization vector itself would imply no "rotation" whatsoever.

We point out that in the general relativistic literature this gravitational polarization rotation is referred to as the *gravitational Faraday effect*, an analogue of the electromagnetic Faraday rotation in a magnetized medium. It was first studied in the problem of gravitomagnetic drag by strong or weak gravitational field near massive rotating bodies.[123–128] The work of Ref.,[58] to the best of our knowledge, is the first to consider this effect in a cosmological context to derive weak lensing distortion to CMB polarization anisotropies.

We are now ready to derive the polarization rotation. We first screen-project ε^μ_e to remove gauge redundancy. Without loss of generality, we may choose the observed polar-

ization (boundary condition here) to be screen-projected with respect to the z direction (which is also the line-of-sight direction so that $\hat{n}_i = \delta_{i3}$), i.e. we choose $\varepsilon_o^{(0)} = 0$ and $\varepsilon_o^{(3)} = 0$. Consistently calculating to linear order in perturbations, we obtain the components of $(S_e \cdot \varepsilon_e)^\mu$,

$$
\begin{aligned}
(S_e \cdot \varepsilon_e)^0 &= \varepsilon_o^{(i)} b_{e,i}, && (4.31)\\
(S_e \cdot \varepsilon_e)^i &= \varepsilon_o^{(i)} - \frac{1}{2}\varepsilon_o^{(j)} (h_e)^i{}_j + \frac{1}{2}\varepsilon_o^{(j)} \int_o^e dr \left[\partial_j b^i - \partial^i b_j - \hat{n}^k \left(\partial_j h^i{}_k - \partial^i h_{jk}\right)\right] && (4.32)\\
&\quad -\frac{1}{2}\hat{n}^i \varepsilon_o^{(j)} \left(b_{e,j} - b_{o,j}\right) + \hat{n}^i \varepsilon_o^{(j)} \int_o^e dz \left[\partial_j a + \frac{1}{2}\dot{b}_j + \frac{1}{2}\hat{n}^k \partial_j b_k - \frac{1}{2}\hat{n}^k \dot{h}_{jk}\right].
\end{aligned}
$$

We then expand the gauge fixed polarization vector $(S_e \cdot \varepsilon_e)^\mu$ using the tetrads at the source location and find the expansion coefficients to be

$$
\begin{aligned}
\varepsilon_e^{(i)} &= \varepsilon_o^{(i)} + \frac{1}{2}\varepsilon_o^{(j)} \int_o^e dr \left[\partial_j b^i - \partial^i b_j - \hat{n}^k \left(\partial_j h^i{}_k - \partial^i h_{jk}\right)\right]\\
&\quad -\frac{1}{2}\hat{n}^i \varepsilon_o^{(j)} \left(b_{e,j} - b_{o,j}\right) + \hat{n}^i \varepsilon_o^{(j)} \int_o^e dr \left[\partial_j a + \frac{1}{2}\dot{b}_j + \frac{1}{2}\hat{n}^k \partial_j b_k - \frac{1}{2}\hat{n}^k \dot{h}_{jk}\right], && (4.33)
\end{aligned}
$$

for $i = 1, 2, 3$. We are not interested in $\varepsilon_e^{(3)}$ as it merely gives the component along the direction of propagation. [2] On the other hand, the two transverse components of our interests, $i = 1, 2$, are found to have undergone a rotation

$$
\varepsilon_e^{(i)} = R^i{}_j \, \varepsilon_o^{(j)}, \qquad (4.34)
$$

[2] Strictly speaking, $e_{e,\mu}^{(3)}$ does not precisely give the direction of propagation at the source location, and $\varepsilon_e^{(3)}$ is not precisely vanishing. However, the difference is essentially due to the deflection of photon momentum, an effect already derived in Section 4.1.1. For a quantification of polarization rotation at leading non-zero order, it is sufficient to neglect this subtlety.

with the rotational matrix given by

$$R^i{}_j = \Pi^i{}_j + \frac{1}{2}\Pi^i{}_{i'}\Pi^{j'}{}_j \int_o^e dr \left[\partial_{j'} b^{i'} - \partial^{i'} b_{j'} - \hat{n}^k \left(\partial_{j'} h^{i'}{}_k - \partial^{i'} h_{j'k}\right)\right]. \quad (4.35)$$

This parametrizes a two-dimensional rotation of an angle ψ (when viewed by the observer against the direction of propagation, counter-clockwise rotation has positive ψ)

$$R^i{}_j = \Pi^i{}_j + \psi\, \epsilon^i{}_{lj'} \hat{n}^l \Pi^{j'}{}_j, \quad (4.36)$$

where ψ up to first order in perturbation is given by

$$\psi = \frac{1}{2}\epsilon^{kij}\hat{n}_k \int_o^e dr \left(\partial_i b_j - \hat{n}^l \partial_i h_{jl}\right). \quad (4.37)$$

For scalar gravitational potentials $a = \Phi$, $b_i = 0$, $h_{ij} = -2\Psi\,\delta_{ij}$, one can verify that the rotation angle ψ vanishes. A non-zero rotation angle may arise from a vector metric perturbation, either from the transverse part of b_i, or from the vectorial part of h_{ij}; it may also arise from gravitational waves, which are parametrized by the transverse trace-free part of h_{ij} (c.f. SVT decomposition Eq. (2.2)).

Eq. (4.37) is our central result for gravitational polarization rotation. We emphasize that the line of sight integral is not a gauge artifact. In fact, under a general gauge transformation Eq. (A.9) (see discussion in Appendix A), Eq. (4.37) shifts only by a boundary

term,

$$\begin{aligned}\psi &\rightarrow \psi + \frac{1}{2}\epsilon^{kij}\,\hat{n}_k \int_o^e dr \left[\partial_i\left(\dot{\xi}_j - \partial_j\xi\right) - \hat{n}^l\partial_i\left(\partial_j\xi_l + \partial_l\xi_j\right)\right] \\ &= \psi - \frac{1}{2}\epsilon^{kij}\,\hat{n}_k \int_o^e dr \frac{d(\partial_i\xi_j)}{dr} \\ &= \psi - \frac{1}{2}\epsilon^{kij}\hat{n}_k\left[(\partial_i\xi_j)_e - (\partial_i\xi_j)_o\right]. \end{aligned} \tag{4.38}$$

This boundary term is understood as a rotation of the reference tetrads under Eq. (A.9),

$$R_{\xi,ij} = \delta_{ij} + \Omega_{ij} = \delta_{ij} + \frac{1}{2}\left(\partial_i\xi_j - \partial_j\xi_i\right), \tag{4.39}$$

according to Eq. (A.11). Of course, the orientation of the physical photon polarization is unaffected, but merely its expansion coefficients with respect to the chosen tetrad vectors change. Therefore, up to the boundary term in Eq. (4.38), the rotation angle ψ is physical. Note that for scalar metric perturbation, ξ^i is a pure gradient field, so that this boundary term vanishes identically. On the other hand, tensor metric perturbations are gauge invariant,[129] for which the boundary term does not exist. These are the two cases most relevant to applications in standard cosmology.

4.1.3 Expanding FLRW universe

In the cosmological context, photons do propagate over large spatial and temporal scales over which the expansion of the Universe is very significant. Therefore, hereafter we generalize our results to an expanding FLRW universe perturbed by metric fluctuations. For

an expanding FLRW universe, and in a general gauge, the perturbed metric is dependent on the conformal time η and the comoving position x^i

$$ds^2 = a^2(\tau)\left[-(1+2A(\mathbf{x},\tau))d\tau^2 + 2B_i(\mathbf{x},\tau)d\tau\, dx^i + (\delta_{ij}+h_{ij}(\mathbf{x},\tau))\, dx^i dx^j\right], \quad (4.40)$$

where τ is the conformal time variable, $a(\tau)$ is the scale factor, and $\mathbf{x}$ is the *comoving* position. Again, the metric perturbation consists of a scalar field $A(\mathbf{x},\tau)$, a vector $B_i(\mathbf{x},\tau)$, and a symmetric tensor $h_{ij}(\mathbf{x},\tau)$. This is conformally similar to the metric in Eq. (4.1).

Photon geodesics are invariant under this Weyl transformation, and hence all angles are invariant. For this reason, the physical effects of metric perturbations on photons in a perturbed FLRW spacetime can be directly obtained from the case of a perturbed Minkowskian spacetime.[121] First, the effect of scale factor $a(\tau)$ can be accounted for by a homogeneous rescaling of the photon 4-momentum $p^\mu \to p^\mu/a^2(\tau)$ and a rescaling of the tetrads $e^\mu_{(a)} \to e^\mu_{(a)}/a(\tau)$. The radial interval dr has to be replaced by the *comoving* radial interval $d\chi = dr/a(\tau)$. Using these rules, results for an expanding Universe can be straightforwardly written down. The fractional anomalous shift in the photon energy, a generalization of Eq. (4.13), reads

$$\frac{a_e\,\omega_e}{a_o\,\omega_o} - 1 = A_e - A_o + \hat{n}^j\,(B_{e,i} - B_{o,i}) + \int_o^e d\chi\left[A' - 2n^i\partial_i A - \hat{n}^i\hat{n}^j\partial_i B_j + \frac{1}{2}n^i n^j h'_{ij}\right]. \quad (4.41)$$

We have used a prime to denote a derivative with respect to the conformal time τ. The

result for transverse angular deflection, a generalization of Eq. (4.18), reads

$$\begin{aligned}\theta^i &= \Pi^i{}_j \left\{ \left(-B_o^j + \frac{1}{2}\hat{n}^k (h_o)^j{}_k \right) \right. \\ &\left. + \int_o^e d\chi \left[\left(1 - \frac{\chi}{\chi_e} \right) \left(-\partial^j A - \hat{n}^k \partial_j B_k + \frac{1}{2}\hat{n}^k \hat{n}^l \partial^j h_{kl} \right) + \frac{1}{\chi_e} \left(B^j - \hat{n}^k h^j{}_k \right) \right] \right\}. \end{aligned} \tag{4.42}$$

Finally, the polarization rotation angle Eq. (4.37) is generalized to

$$\psi = \frac{1}{2}\epsilon^{kij}\, \hat{n}_k \int_o^e d\chi \left(\partial_i B_j - \hat{n}^l \partial_i h_{jl} \right). \tag{4.43}$$

Again, all line-of-sight integrals are understood to have their integrands evaluated on the past light cone $d/d\chi = -\partial_\tau + \hat{n}^i \partial_i$.

4.2 Lensing kernels in harmonic space

When we observe the photons traveling in all directions from the last scattering surface, different line of sights probe different intervening metric fluctuation. Those are random but correlated to some extent. The motivation of an all-sky harmonic analysis is to disentangle, for any physical observable depending on the direction in the sky $\hat{\mathbf{n}}$, variations that occur across different angular scales (inversely proportional to the harmonic multipole l).

In this Section, we provide a unified framework to analyze the effect of weak lensing (both deflection of line of sight and rotation of polarization) on temperature and polarization anisotropies. We will see that the effect of weak lensing, to leading order in perturbation, is to couple together different temperature or polarization spherical harmonic multipoles,

which are presumably Gaussian and uncorrelated with each other in an unlensed CMB sky.

Our discussion will be based on the TAM formalism introduced in Chapter 2. We will borrow some of the notation adopted there.

4.2.1 Deflection potentials

Deflection is described by a displacement field $\theta^i(\hat{\mathbf{n}})$ on the two-dimensional sky. A general displacement field may be decomposed into a curl-free part and a divergence-free part, similar to the transverse/longitudinal decomposition of a three-dimensional vector field (c.f. Eq. (2.1)). Therefore, the displacement is generated by two scalar potentials: a gradient deflection potential $\phi(\hat{\mathbf{n}})$ and a curl-deflection potential $\Omega(\hat{\mathbf{n}})$.[86,90] We have

$$\theta^i(\hat{\mathbf{n}}) \;=\; -M_\perp^i\,\phi(\hat{\mathbf{n}}) + K^i\,\Omega(\hat{\mathbf{n}}), \tag{4.44}$$

where differential operators K^i and $M_\perp^i$ have been defined in Eq. (2.28). Both potentials admit a spherical harmonic expansion on the sky,

$$\phi(\hat{\mathbf{n}}) = \sum_{JM} \phi_{JM} Y_{(JM)}(\hat{\mathbf{n}}), \qquad \Omega(\hat{\mathbf{n}}) = \sum_{JM} \Omega_{JM} Y_{(JM)}(\hat{\mathbf{n}}). \tag{4.45}$$

If the foreground metric perturbations are statistically isotropic, the potentials have angular power spectra

$$\langle \phi_{JM} \phi^*_{J'M'} \rangle = C_J^{\phi\phi} \delta_{JJ'} \delta_{MM'}, \qquad \langle \Omega_{JM} \Omega^*_{J'M'} \rangle = C_J^{\Omega\Omega} \delta_{JJ'} \delta_{MM'}, \tag{4.46}$$

where we average over all possible realizations. Note that the statistics of metric perturbations are not necessarily homogeneous; since only perturbations on the past light cone are relevant, they may evolve significantly along the radial direction. The deflection field then may be alternatively expanded using vector spherical harmonics

$$\theta^i(\hat{\mathbf{n}}) \;=\; \sum_{JM} \sqrt{J(J+1)} \left(-\phi_{JM} Y^{E,i}_{(JM)}(\hat{\mathbf{n}}) + \Omega_{JM} Y^{B,i}_{(JM)}(\hat{\mathbf{n}}) \right). \tag{4.47}$$

We see that gradient-deflection ϕ generates parity-even displacement patterns, while curl-deflection Ω generates parity-odd displacement patterns.

4.2.2 Deflection of temperature field

The *unlensed* temperature fluctuation field $\Theta(\hat{\mathbf{n}})$, being a scalar function, can be expanded using the usual spherical harmonics [3]

$$\Theta(\hat{\mathbf{n}}) \;=\; \sum_{lm} \Theta_{lm}\, Y_{(lm)}(\hat{\mathbf{n}}). \tag{4.48}$$

The temperature multipoles Θ_{lm} are intrinsically Gaussian random variables with isotropic statistics. The complete statistical property is quantified by the unlensed power spectrum

$$\langle \Theta_{lm} \Theta^*_{l'm'} \rangle \;=\; C_l^{\Theta\Theta} \delta_{ll'} \delta_{mm'}. \tag{4.49}$$

[3] A notational remark. We reserve uppercase J, M for multipoles describing lens scales, and lowercase l, m for multipoles describing CMB fluctuation scales.

Since deflection moves lines of sight, but otherwise preserves surface brightness, as dictated by Liouville's theorem, the lensed temperature field, always labelled with a tilde, is given by a displaced one[130]

$$\tilde{\Theta}(\hat{\mathbf{n}}) \;=\; \Theta(\hat{\mathbf{n}}) + \theta^i\,\nabla_i\Theta(\hat{\mathbf{n}}) + \frac{1}{2}\theta^i\theta^j\,\nabla_i\nabla_j\Theta(\hat{\mathbf{n}}) + \cdots\,, \tag{4.50}$$

where ∇_i is the covariant derivative defined on the two-sphere [4].

Note that we have truncated the Taylor expansion at quadratic order, which will be sufficient for deriving the leading-order weak lensing effect on CMB power spectra. The caveat is that on sufficiently small angular scales, say below the typical deflection angle, such a perturbative expansion fails to converge, and non-perturbative evaluation is in need.[131] The interpretation for perturbative expansion Eq. (4.50) with non-commuting covariant derivative is that the line of sight is being shifted along a big circle. This defines the deflection beyond the leading-order, flat-sky approximation, and carefully accounts for the curvature of the sky at higher orders in the expansion. Note that a careful definition of deflection on a spherical sky is crucial already at the leading-order computation of lensed power spectrum. The inclusion of the quadratic terms in Eq. (4.50) guarantees that the ensemble averaged magnification on the source plane is unity, a result due to overall conservation of light rays, for both weak or strong gravitational lensing.[132] As we will see later, this issue is even more subtle when deflection of the polarization field is being considered.

We also introduce here a rigorous formulation of deflection beyond linear order in deriva-

[4]When *acting on scalar functions* $\nabla_i = -M_{\perp i}$; when acting on vector functions or functions with more indices, the effect of ∇_i is more complex.

tive. For every direction in the sky $\hat{n}^i$, there is a unique displacement θ^i, whose effect, at line of sight $\hat{n}^i$ and to nonlinear orders on a spherical sky, is equivalent to a rotation of geodesic distance $|\theta^i|$ about the direction $a^i(\hat{\mathbf{n}}) = \epsilon^{ijk}\hat{n}_j\theta_k(\hat{\mathbf{n}})$ perpendicular to both θ_i and $\hat{n}_i$. Since the generator of such a rotation is the orbital angular momentum operator L_i, the deflected field reads

$$\tilde{\Theta}(\hat{\mathbf{n}}) = \mathcal{N}\left[\exp\left(ia^i(\hat{\mathbf{n}})L_i\right)\right]\Theta(\hat{\mathbf{n}}) = \mathcal{N}\left[\exp\left(-i\theta^i(\hat{\mathbf{n}})M_{\perp i}\right)\right]\Theta(\hat{\mathbf{n}}). \tag{4.51}$$

Here $\mathcal{N}[\cdots]$ is *normal ordering*, i.e. in the Taylor expansion of the exponential all θ^i's should be positioned to the left of all differential operators,

$$\mathcal{N}\left[\exp\left(-i\theta^i M_{\perp i}\right)\right] = \sum_{n=0}^{\infty}\frac{(-)^n}{n!}\theta^{i_1}\theta^{i_2}\cdots\theta^{i_n}M_{\perp i_1}M_{\perp i_2}\cdots M_{\perp i_n}. \tag{4.52}$$

This prescription exactly leads to Eq. (4.50).

Having the derivative expansion Eq. (4.50), the lensed temperature multipoles are then given by

$$\tilde{\Theta}_{lm} \;=\; \Theta_{lm} + \int d^2n\,[Y_{(lm)}]^*\,\theta^i\nabla_i\Theta + \frac{1}{2}\int d^2n\,[Y_{(lm)}]^*\,\theta^i\theta^j\nabla_i\nabla_j\Theta. \tag{4.53}$$

We can then combine Eq. (4.53) with expansion Eq. (4.47) to compute lensed power spectra $\left\langle\tilde{\Theta}_{lm}\tilde{\Theta}^*_{l'm'}\right\rangle = \tilde{C}_l^{\Theta\Theta}\delta_{ll'}\delta_{mm'}$. Here the ensemble average is carried out over all realizations of the intrinsic CMB fluctuation *and* all realizations of the metric lens. It is normally expected that the power spectrum remains diagonal and isotropic if both ensemble averages are taken.

We furthemore assume that intrinsic CMB fluctuations Θ are *uncorrelated* with the lensing potentials ϕ and Ω, so that ensemble average over intrinsic CMB realizations and over lens realizations decouple. This assumption is valid for lensing by foreground large-scale structure mainly at low redshifts $z \sim 3$, which is spatially far away from the last scattering surface. The assumption would also be valid for lensing by a stochastic background of primordial gravitational waves, which is expected to be statistically independent of the primordial adiabatic perturbations that serve as the seeds of cosmic structure. Some useful algebraic results crucial for our derivation are collected in Appendix B. The end result for correction $\delta C_l \equiv \tilde{C}_l - C_l$ reads

$$\begin{aligned}\delta C_l^{\Theta\Theta} &= -l(l+1)R\,C_l^{\Theta\Theta} + \sum_{l'} C_{l'}^{\Theta\Theta} l'(l'+1)(2l'+1) \\ &\times \left[\sum_J J(J+1)\frac{2J+1}{4\pi}\begin{pmatrix} l_1 & J_1 & l \\ 1 & -1 & 0 \end{pmatrix}^2 \left(P^+_{l'Jl}C_J^{\phi\phi} + P^-_{l'Jl}C_J^{\Omega\Omega}\right)\right], \quad (4.54)\end{aligned}$$

where

$$R \equiv \frac{1}{2}\sum_J J(J+1)\frac{2J+1}{4\pi}\left(C_J^{\phi\phi} + C_J^{\Omega\Omega}\right) \quad (4.55)$$

is the all-sky *mean square* deflection angle, and the parity symbol is defined as $P^{\pm}_{l_1 l_2 l_3} \equiv [1 \pm (-)^{l_1+l_2+l_3}]/2$.

4.2.3 Deflection of polarization field

Photon polarization is quantified by a symmetric, trace-free rank-two tensor $P_{ij}(\hat{n})$ perpendicular to the line of sight. When circular polarization is neglected, it contains two degrees of freedom, corresponding to the two conventional Stokes parameters Q and U. It admits an expansion, analogous to Eq. (4.48), in terms of tensor spherical harmonics (c.f. Eq. (2.59))

$$P_{ij}(\hat{\mathbf{n}}) = \sum_{lm} \left(E_{lm}\, Y^{TE}_{(lm)ij}(\hat{\mathbf{n}}) + B_{lm}\, Y^{TB}_{(lm)ij}(\hat{\mathbf{n}}) \right), \tag{4.56}$$

where the E-modes E_{lm}'s correspond to parity-even patterns and the B-modes B_{lm}'s correspond to parity-odd patterns. The unlensed polarization field is Gaussian random, quantified by the auto/cross power spectra

$$\langle E_{lm} E^*_{l'm'} \rangle = C_l^{EE} \delta_{ll'} \delta_{mm'}, \tag{4.57}$$

$$\langle B_{lm} B^*_{l'm'} \rangle = C_l^{BB} \delta_{ll'} \delta_{mm'}, \tag{4.58}$$

$$\langle \Theta_{lm} E^*_{l'm'} \rangle = C_l^{\Theta E} \delta_{ll'} \delta_{mm'}, \tag{4.59}$$

$$\langle E_{lm} B^*_{l'm'} \rangle = 0, \tag{4.60}$$

$$\langle \Theta_{lm} B^*_{l'm'} \rangle = 0. \tag{4.61}$$

Cross power spectra between the B modes with either the temperature or the E modes vanish due to invariant statistics under a parity flip.

Similar to the case of temperature, deflection is equivalent to a parallel transport of

$P_{ij}(\hat{\mathbf{n}})$ along the great circle over a geodesic distance of $|\theta^i|$, which is generated by a rotation along the direction $a^i(\hat{\mathbf{n}}) = \epsilon^{ijk}\hat{n}_j\theta_k(\hat{\mathbf{n}})$. What is different, though, is that it is the total angular momentum operator J_i, rather than the orbital angular momentum operator L_i, that is the appropriate generator for spin-two polarization field. We have

$$\tilde{P}_{ij}(\hat{\mathbf{n}}) = \mathcal{N}\left[\exp\left(-i\epsilon^{klm}\hat{n}_l\theta_m(\hat{\mathbf{n}})J_k\right)\right]P_{ij}(\hat{\mathbf{n}}). \tag{4.62}$$

For calculating power spectra at leading non-trivial order, it suffices to expand to quadratic order in derivatives,

$$\begin{aligned}\tilde{P}_{ij}(\hat{\mathbf{n}}) &= P_{ij}(\hat{\mathbf{n}}) - i\theta^k(\hat{\mathbf{n}})\epsilon_{klm}\hat{n}_l J_m P_{ij}(\hat{\mathbf{n}}) \\ &\quad -\frac{1}{2}\theta^k(\hat{\mathbf{n}})\theta^{k'}(\hat{\mathbf{n}})(\epsilon_{klm}\hat{n}_l J_m)(\epsilon^{k'l'm'}\hat{n}_{l'}J_{m'})P_{ij}(\hat{\mathbf{n}}) + \cdots\end{aligned} \tag{4.63}$$

Using Eq. (4.56), the polarization multipole moments transform under deflection as

$$\begin{aligned}\tilde{E}_{lm} &= E_{lm} - i\int d^2n\,\left[Y^{TE}_{(lm)i''j''}\right]^*\theta^i\epsilon_{ijk}\hat{n}^j J^k P_{i''j''} \\ &\quad -\frac{1}{2}\int d^2n\,\left[Y^{TE}_{(lm)i''j''}\right]^*\theta^i\theta^{i'}(\epsilon_{ijk}\hat{n}^j J^k)(\epsilon_{i'j'k'}\hat{n}^{j'}J^{k'})P_{i''j''}, \qquad (4.64) \\ \tilde{B}_{lm} &= B_{lm} - i\int d^2n\,\left[Y^{TB}_{(lm)i''j''}\right]^*\theta^i\epsilon_{ijk}\hat{n}^j J^k P_{i''j''} \\ &\quad -\frac{1}{2}\int d^2n\,\left[Y^{TB}_{(lm)i''j''}\right]^*\theta^i\theta^{i'}(\epsilon_{ijk}\hat{n}^j J^k)(\epsilon_{i'j'k'}n^{j'}J^{k'})P_{i''j''}. \qquad (4.65)\end{aligned}$$

We will see later that Eqs. (4.64)–(4.65) are *not* the complete results for the lensed polarization; gravitational Faraday rotation has to be included in addition to deflection.

4.2.4 Polarization rotation

So far we have only considered one effect from the metric perturbations on photon polarization, namely the deflection $\theta^i(\hat{\mathbf{n}})$ of light of sight. Another effect, gravitational Faraday rotation $\psi(\hat{\mathbf{n}})$, has to be simultaneously accounted for. The reason is that polarization rotation field ψ is in full correlation with the curl-deflection potential Ω, as one can use Eq. (4.42), Eq. (4.43) and Eq. (4.44) to directly verify that for *any lens realization*

$$\psi(\hat{\mathbf{n}}) \;=\; -\frac{1}{2}K^i\,\theta_i(\hat{\mathbf{n}}) = -\frac{1}{2}K^i K_i\,\Omega(\hat{\mathbf{n}}). \tag{4.66}$$

Since the curl-deflection potential Ω generates the anti-symmetric part of the lensing Jacobian, it quantifies rotation of an infinitesimal image with respect to its source orientation. The above equation relating ψ and Ω states that *the polarization rotation is exactly equal to the image rotation in weak gravitational lensing.* With a spherical harmonic expansion of ψ,

$$\psi(\hat{\mathbf{n}}) \;=\; \sum_{JM}\psi_{JM}Y_{(JM)}(\hat{\mathbf{n}}), \tag{4.67}$$

that implies relations between the rotation power spectrum, the curl-deflection power spectrum, and their cross power spectrum,

$$C_J^{\psi\psi} = \frac{J(J+1)}{2}C_J^{\Omega\psi} = \frac{J^2(J+1)^2}{4}C_J^{\Omega\Omega}. \tag{4.68}$$

We are now in a position to derive the additional change in the polarization tensor $P_{ij}(\hat{\mathbf{n}})$

due to rotation ψ. To derive at sufficient order, the linearized rotation matrix Eq. (4.36) may be exponentiated and then be truncated at quadratic order

$$R^i{}_j = \left(\delta^i{}_{i'} + \psi\,\epsilon^i{}_{li'}\hat{n}^l + \frac{1}{2}\psi^2\,\epsilon^i{}_{li''}\hat{n}^l\epsilon^{i''ki'}\hat{n}_k\right)\Pi_{i'j} + \cdots \tag{4.69}$$

The polarization tensor transforms as

$$\begin{aligned}\tilde{P}_{ij} &= R_i{}^{i'}R_j{}^{j'}P_{i'j'} \\ &= P_{ij} + 2\psi\,\epsilon_{ik}{}^{i'}\hat{n}^k P_{i'j} + \psi^2\left(\epsilon_{ili''}\hat{n}^l\epsilon^{i''ki'}\hat{n}_k P_{i'j} + \epsilon_{il}{}^{i'}\hat{n}^l\epsilon_j{}^{kj'}\hat{n}_k P_{i'j'}\right).\end{aligned} \tag{4.70}$$

By definition, the helicity basis tensor spherical harmonics $Y^{\pm 2}_{(lm)ij}$, defined in the text before Eq. (2.69), pick up a phase under this rotation

$$\begin{aligned}R_i{}^{i'}R_j{}^{j'}Y^{\pm 2}_{(l'm')i'j'}(\hat{\mathbf{n}}) &= \exp\left[\mp 2i\psi(\hat{\mathbf{n}})\right]Y^{\pm 2}_{(l'm')ij}(\hat{\mathbf{n}}) \\ &= \left[1 \mp 2i\psi(\hat{\mathbf{n}}) - 2\psi^2(\hat{\mathbf{n}})\right]Y^{\pm 2}_{(l'm')ij}(\hat{\mathbf{n}}).\end{aligned} \tag{4.71}$$

Since the TE/TB type tensor spherical harmonics $Y^{TE,TB}_{(lm)ij}$ are linear combinations of the helicity harmonics $Y^{\pm 2}_{(lm)ij}$, we find lensing corrections to the polarization multipole moments

(in additional to Eqs. (4.64)–(4.65))

$$\begin{aligned}
\tilde{E}_{lm} - E_{lm} \quad \supset \quad & 2\sum_{l'm'}\sum_{JM}\psi_{JM}\left(E_{l'm'}\int d^2n\,\left[Y^{TE,ij}_{(lm)}\right]^*Y_{(JM)}Y^{TB}_{(l'm')ij}\right. \\
& \left.-B_{l'm'}\int d^2n\,\left[Y^{TE,ij}_{(lm)}\right]^*Y_{(JM)}Y^{TE}_{(l'm')ij}\right) \\
& -2\sum_{l'm'}\sum_{JM}\sum_{J'M'}\psi^*_{J'M'}\psi_{JM}\left(E_{l'm'}\int d^2n\,\left[Y^{TE,ij}_{(lm)}\right]^*[Y_{(J'M')}]^*Y_{(JM)}Y^{TE}_{(l'm')ij}\right. \\
& \left.+B_{l'm'}\int d^2n\,\left[Y^{TE,ij}_{(lm)}\right]^*[Y_{(J'M')}]^*Y_{(JM)}Y^{TB}_{(l'm')ij}\right), \qquad (4.72) \\
\tilde{B}_{lm} - B_{lm} \quad \supset \quad & 2\sum_{l'm'}\sum_{JM}\psi_{JM}\left(E_{l'm'}\int d^2n\,\left[Y^{TB,ij}_{(lm)}\right]^*Y_{(JM)}Y^{TB}_{(l'm')ij}\right. \\
& \left.-B_{l'm'}\int d^2n\,\left[Y^{TB,ij}_{(lm)}\right]^*Y_{(JM)}Y^{TE}_{(l'm')ij}\right) \\
& -2\sum_{l'm'}\sum_{JM}\sum_{J'M'}\psi^*_{J'M'}\psi_{JM}\left(E_{l'm'}\int d^2n\,\left[Y^{TB,ij}_{(lm)}\right]^*[Y_{(J'M')}]^*Y_{(JM)}Y^{TE}_{(l'm')ij}\right. \\
& \left.+B_{l'm'}\int d^2n\,\left[Y^{TB,ij}_{(lm)}\right]^*[Y_{(J'M')}]^*Y_{(JM)}Y^{TB}_{(l'm')ij}\right). \qquad (4.73)
\end{aligned}$$

Note that when computing lensed power spectra, contributions from rotation Eqs. (4.72)–(4.73) must be combined with contributions from deflection, namely Eq. (4.50) and Eqs. (4.64)–(4.65), due to correlation between Ω and ψ.

After some heavy but straightforward algebra, we report the final results for the lensed temperature/polarization (cross-)power spectra, taking into account deflection of both type

ϕ and Ω, as well as the rotation ψ:

$$
\begin{aligned}
\delta C_l^{\Theta\Theta} &= -l\left(l+1\right) R C_l^{\Theta\Theta} \\
&\quad + \frac{1}{2l+1}\sum_{l'J}\left[C_J^{\phi\phi}\left(F^{\phi}_{ll'J}\right)^2 P^{+}_{ll'J} C_{l'}^{\Theta\Theta} + C_J^{\Omega\Omega}\left(F^{\Omega}_{ll'J}\right)^2 P^{-}_{ll'J} C_{l'}^{\Theta\Theta}\right], \\
\delta C_l^{EE} &= -\left(l^2+l-4\right) R C_l^{EE} - 4 S C_l^{EE} \\
&\quad + \frac{1}{2l+1}\sum_{l'J}\left[C_J^{\phi\phi}\left(G^{\phi}_{ll'J}\right)^2 \left(P^{+}_{ll'J} C_{l'}^{EE} + P^{-}_{ll'J} C_{l'}^{BB}\right) + C_J^{\Omega\Omega}\left(G^{\Omega}_{ll'J}\right)^2 \left(P^{-}_{ll'J} C_{l'}^{EE} + P^{+}_{ll'J} C_{l'}^{BB}\right)\right] \\
&\quad + \frac{4}{2l+1}\sum_{l'J} C_J^{\psi\psi}\left(H^{\psi}_{ll'J}\right)^2 \left(P^{-}_{ll'J} C_{l'}^{EE} + P^{+}_{ll'J} C_{l'}^{BB}\right) \\
&\quad - \frac{4}{2l+1}\sum_{l'J} C_J^{\Omega\psi} G^{\Omega}_{ll'J} H^{\psi}_{ll'J}\left(P^{-}_{ll'J} C_{l'}^{EE} + P^{+}_{ll'J} C_{l'}^{BB}\right), \\
\delta C_l^{BB} &= -\left(l^2+l-4\right) R C_l^{BB} - 4 S C_l^{BB} \\
&\quad + \frac{1}{2l+1}\sum_{l'J}\left[C_J^{\phi\phi}\left(G^{\phi}_{ll'J}\right)^2 \left(P^{-}_{ll'J} C_{l'}^{EE} + P^{+}_{ll'J} C_{l'}^{BB}\right) + C_J^{\Omega\Omega}\left(G^{\Omega}_{ll'J}\right)^2 \left(P^{+}_{ll'J} C_{l'}^{EE} + P^{-}_{ll'J} C_{l'}^{BB}\right)\right] \\
&\quad + \frac{4}{2l+1}\sum_{l'J} C_J^{\psi\psi}\left(H^{\psi}_{ll'J}\right)^2 \left(P^{+}_{ll'J} C_{l'}^{EE} + P^{-}_{ll'J} C_{l'}^{BB}\right) \\
&\quad - \frac{4}{2l+1}\sum_{l'J} C_J^{\Omega\psi} G^{\Omega}_{ll'J} H^{\psi}_{ll'J}\left(P^{+}_{ll'J} C_{l'}^{EE} + P^{-}_{ll'J} C_{l'}^{BB}\right), \\
\delta C_l^{\Theta E} &= -\left(l^2+l-2\right) R C_l^{\Theta E} \\
&\quad - \frac{1}{2l+1}\sum_{l'J}\left[C_J^{\phi\phi} F^{\phi}_{ll'J} G^{\phi}_{ll'J} P^{+}_{ll'J} C_{l'}^{\Theta E} + C_J^{\Omega\Omega} F^{\Omega}_{ll'J} G^{\Omega}_{ll'J} P^{-}_{ll'J} C_{l'}^{\Theta E}\right] \\
&\quad + \frac{2}{2l+1}\sum_{l'J} C_J^{\Omega\psi} F^{\Omega}_{ll'J} H^{\psi}_{ll'J} P^{-}_{ll'J} C_{l'}^{\Theta E}.
\end{aligned}
\tag{4.74}
$$

Here we have introduced five lensing kernels

$$F^{\phi}_{ll'J} = F^{\Omega}_{ll'J} = -\sqrt{J(J+1)l'(l'+1)}\sqrt{\frac{\Pi_{ll'J}}{4\pi}}\begin{pmatrix} l & J & l' \\ 0 & -1 & 1 \end{pmatrix}, \tag{4.75}$$

$$G^{\phi/\Omega}_{ll'J} = \sqrt{\frac{J(J+1)}{2}}\sqrt{\frac{\Pi_{ll'J}}{4\pi}}\left[\sqrt{\frac{(l'+2)(l'-1)}{2}}\begin{pmatrix} l & J & l' \\ 2 & -1 & -1 \end{pmatrix} \pm \sqrt{\frac{(l'+3)(l'-2)}{2}}\begin{pmatrix} l & J & l' \\ 2 & 1 & -3 \end{pmatrix}\right], \tag{4.76}$$

$$H^{\psi}_{ll'J} = \frac{J(J+1)}{2}\sqrt{\frac{\Pi_{ll'J}}{4\pi}}\begin{pmatrix} l & J & l' \\ 2 & 0 & -2 \end{pmatrix}. \tag{4.77}$$

We have also introduced the *mean square* rotation angle

$$S = \sum_J \frac{2J+1}{4\pi} C^{\psi\psi}_J. \tag{4.78}$$

These expressions are the key analytical results of this Chapter. They are valid up to linear order in arbitrary metric perturbation. Although we have referred to the CMB as the most relevant physics context, these results are generally applicable to any wide-angle intensity and polarization surveys analyzed under a spherical harmonic expansion. The deflection kernels $F^{\phi/\Omega}_{ll'J}$ and $G^{\phi/\Omega}_{ll'J}$ have been known in the CMB lensing literature.[90,119,133] The polarization rotation kernel $H^{\psi}_{ll'J}$ has been derived for the first time in our Ref.[58] with a correct treatment of the cross-correlation with deflection. The presence of parity symbols $P^{\pm}_{ll'J}$ is a result of parity conservation under ensemble average. These expressions, at the order we seek, are compatible with *conservation of power* under lensing, i.e. $\sum_l (2l+1)C^{\Theta\Theta}_l$ and $\sum_l (2l+1)[C^{EE}_l + C^{BB}_l]$ are invariant under lensing, for arbitrary intrinsic CMB power

spectra and for arbitrary lens power spectra. This conclusion can be explicitly verified from the analytical expressions we have provided here.

4.3 Lensing by gravitational waves and polarization rotation

As an illustrating example, we compute the weak lensing impact on CMB polarization anisotropies from stochastic gravitational waves, i.e. tensor metric perturbations. This may be a relic of inflation, or is generated by large-scale structure formation or other subhorizon physics. This possibility has been considered previously.[119,133,134] Although current CMB lensing surveys have put tight constraints on this feeble signal,[4] the constraints may improve significantly in the future with high-resolution cosmological 21cm surveys.[87] It has also been claimed in the literature that curl deflection is more efficient in converting E-mode polarization anisotropies from primordial adiabatic perturbations into secondary B-mode polarization anisotropies ("the lensing B modes").[135] Our results falsify this conclusion, attributing the discrepancy to the neglecting of the gravitational Faraday rotation.

4.3.1 Lens power spectra from gravitational waves

We start by considering an FLRW universe perturbed by tensor metric perturbations, which amounts to setting in the general perturbed metric $A = 0$, $B_i = 0$ and $h_{ij} = \gamma_{ij}$, where γ_{ij} is trace-free $\delta^{ij}\gamma_{ij} = 0$ and divergence-free $\partial^i\gamma_{ij} = 0$. The tensor metric perturbation γ_{ij}, being a symmetric tensor field, can be most generally expanded using the TAM basis

waves

$$\gamma_{ij}(\mathbf{x},\tau) \;=\; \int \frac{k^2 dk}{(2\pi)^3} \sum_{JM} \sum_{\alpha=TE,TB} \gamma^{\alpha}_{JM}(k) T_{\gamma}(k,\tau) \left(4\pi i^J\right) \Psi^{k,\alpha}_{(JM)ij}(\mathbf{x}). \tag{4.79}$$

The initial TAM amplitude $\gamma^{\alpha}_{JM}(k)$ is linearly extrapolated to late times by a tensor transfer function $T_{\gamma}(k,\tau)$, which can be solved from the equation of motion for tensor perturbations in an expanding universe

$$\gamma''_{ij} + 2aH\gamma'_{ij} - \partial^2\gamma_{ij} \;=\; 0. \tag{4.80}$$

For example, in a matter-dominated Universe we have $T_{\gamma}(k,\tau) = 3j_1(k\tau)/(k\tau)$. We may then quantify the two-point statistics of a stochastic background using the power spectra in TAM basis

$$\left\langle \gamma^{\alpha}_{JM}(k) \left[\gamma^{\alpha'}_{J'M'}(k')\right]^* \right\rangle = \frac{(2\pi)^3}{k^2} \frac{P_{\gamma}(k)}{2} \delta_D(k-k') \delta_{\alpha\alpha'} \delta_{JJ'} \delta_{MM'}, \quad \alpha = TE,\, TB, \tag{4.81}$$

if the gravitational-wave background is statistically homogeneous and isotropic.

To obtain the relevant lensing spectra induced by gravitational waves, we specialize to the case of tensor metric perturbations in Eq. (4.42) and Eq. (4.43). In this way, we derive

deflection potentials and polarization rotation angle

$$\phi(\hat{\mathbf{n}}) = \sum_{JM} Y_{(JM)}(\hat{\mathbf{n}}) \int \frac{k^2 dk}{(2\pi)^3} \gamma^{TE}_{JM}(k)(4\pi i^J) F^{GW,\phi}_J, \tag{4.82}$$

$$\Omega(\hat{\mathbf{n}}) = \sum_{JM} Y_{(JM)}(\hat{\mathbf{n}}) \int \frac{k^2 dk}{(2\pi)^3} \gamma^{TB}_{JM}(k)(4\pi i^J) F^{GW,\Omega}_J, \tag{4.83}$$

$$\psi(\hat{\mathbf{n}}) = \sum_{JM} \frac{J(J+1)}{2} Y_{(JM)}(\hat{\mathbf{n}}) \int \frac{k^2 dk}{(2\pi)^3} \gamma^{TB}_{JM}(k)(4\pi i^J) F^{GW,\psi}_J. \tag{4.84}$$

It can be seen that the TE-type TAM components only generate gradient-type deflection, while only the TB-type TAM components are responsible for curl-type deflection, as well as polarization rotation. The relevant radial kernels are given by[42,121]

$$F^{GW,\phi}_J(k) = \sqrt{\frac{(J-1)(J+2)}{2J(J+1)}} \int_0^{k\chi} d(k\chi') \frac{T_\gamma(k, \tau_0 - \chi')}{k\chi'^2} \left\{ \left[j'_J\left(k\chi'\right) + \frac{j_J\left(k\chi'\right)}{k\chi'} \right] - \frac{J(J+1)}{2} \frac{\chi - \chi'}{\chi} \frac{j_J\left(k\chi'\right)}{k\chi'} \right\}, \tag{4.85}$$

$$F^{GW,\Omega}_J(k) = F^{GW,\psi}_J(k) = i\sqrt{\frac{(J-1)(J+2)}{2J(J+1)}} \frac{1}{\chi} \int_0^{k\chi} d(k\chi')\, T_\gamma(k, \tau_0 - \chi') \frac{j_J(k\chi')}{(k\chi')^2}. \tag{4.86}$$

Here χ is the *comoving* radial distance to the source plane, and τ_0 is the present conformal time. Note that the same kernel appears for the curl-deflection Ω (equivalent to image rotation) and the polarization rotation ψ. The two-dimensional lens power spectra then follow directly from the three-dimensional gravitational-wave power spectrum Eq. (4.81),

$$C^{\phi\phi}_J = \frac{2}{\pi} \int k^2 dk \frac{P_\gamma(k)}{2} |F^{GW,\phi}_J(k)|^2, \tag{4.87}$$

$$C^{\Omega\Omega}_J = \frac{2}{\pi} \int k^2 dk \frac{P_\gamma(k)}{2} |F^{GW,\Omega}_J(k)|^2, \tag{4.88}$$

$$C^{\psi\psi}_J = \frac{J^2(J+1)^2}{4} \frac{2}{\pi} \int k^2 dk \frac{P_\gamma(k)}{2} |F^{GW,\Omega}_J(k)|^2. \tag{4.89}$$

Supplied with the general results Eq. (4.74), these give predictions for the ensemble-averaged CMB power spectra under weak lensing distortion, to which we turn next.

4.3.2 Corrections to CMB power spectra

We numerically evaluate lensing corrections to the CMB power spectra for the *WMAP*+BAO+H_0 best-fit cosmological parameters.[88] An almost scale-invariant primordial background of inflationary gravitational waves with tensor-to-scalar ratio $r = 0.13$ (the latest experimental constraint $r < 0.12$[34] has not been significantly improved) is considered.

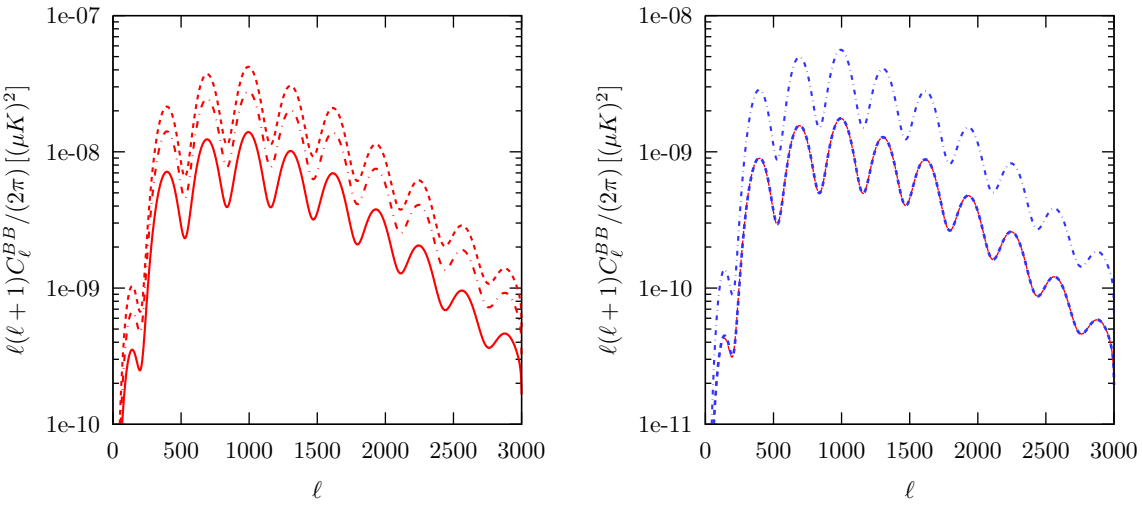


Figure 4.2: *Left panel:* Lensing B-mode power spectrum from an inflationary tensor-mode background $r = 0.13$ in concordance cosmology. B-mode conversion from deflection alone (dashed), from rotation alone (dash-dotted), and from a correlation between deflection and rotation (solid) are compared. *Right panel:* A comparison of how efficient primordial E-mode polarization can be converted into secondary B-mode polarization, between two fictitious universes of pure gradient deflection ϕ and pure curl deflection respectively Ω, but with the same deflection power spectrum numerically $C_J^{\Omega\Omega} = C_J^{\phi\phi}$. We consider three cases: by gradient deflection ϕ only (red, thin dashed), by curl deflection Ω only (blue dash-dotted), and by a full correlation between curl deflection Ω and rotation ψ (blue, thick dashed).

Primordial gravitational waves are expected to generate first-order B-mode polariza-

tions that appear on large angular scales $l \leq 200$. However, toward smaller angular scales, the first-order B-mode signal is expected to be cut off due to rapid dilution of the gravitational waves in the subhorizon regime with the cosmic expansion. Instead, secondary B-modes converted from first-order E modes by weak gravitational lensing dominate on those scales. Apart from the scalar gravitational potential perturbations from foreground large-scale matter distribution, freely propagating gravitational waves act as an additional lens. In Figure 4.2, we focus on the generation of secondary B modes by gravitational waves. It can be seen from the left panel that including deflection but neglecting rotation overestimates the generated power. In fact, by including rotation, the expected power is reduced, due to partial cancellation between the curl deflection and polarization rotation.

The origin of this cancellation is vividly demonstrated in the heuristic picture of Figure 4.3. Suppose we start with polarization anisotropies of purely E-mode type on the last scattering surface, as anticipated in the standard inflation scenario, a foreground curl deflection potential Ω, assumed to be uniform across a small patch of the sky, deflects light rays and hence induces secondary B modes. This is because the optical consequence of such a uniform curl deflection potential is to "rotate" the small patch by some angle, but otherwise to keep the polarization at each point unchanged, which results in a "swirling" pattern. However, this swirling pattern is to some extent spurious, as the same tensor metric perturbations generating this curl deflection potential Ω would also induce polarization rotation ψ, which according to our result Eq. (4.66), is equal to the image rotation. Rotating the polarization at each point by the same angle as that for the image rotation largely undoes the swirling pattern, and therefore greatly reduces the B-mode anisotropy.

It is worth pointing out that the cancellation will be *exact* if curl deflection Ω is perfectly uniform. In more realistic cases, the cancellation will not be complete, since curl deflection Ω generated by a stochastic foreground of any metric perturbations is always spatially varying.

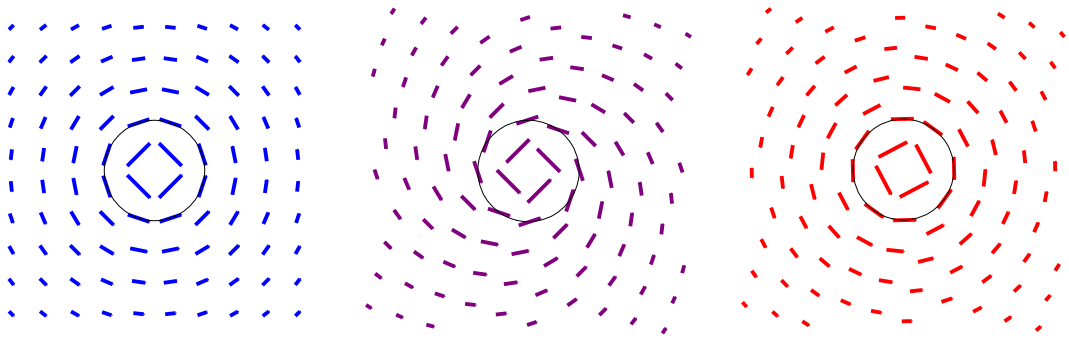


Figure 4.3: A polarization map with only E modes on the last scattering surface (left) may be deflected by a foreground of uniform curl potential Ω to give rise to secondary B modes (middle). As indicated by the black circle, this curl deflection deflects lines of sight along the tangential direction (clockwise) of concentric circles. However, when polarization rotation is simultaneously included, the B modes are reduced (right).

It was previously found in Ref.[135] that compared to the gradient deflection ϕ, the curl deflection Ω can convert primary E modes into secondary B modes with greater efficiency. However, the polarization rotation ψ, which is expected to be (anti-)correlated with the curl deflection, was not accounted for at the time of their work. The right panel of Figure 4.2 clearly shows that the efficiency is almost identical (difference is not discernible from the plot on almost all angular scales) when ψ is accounted for, at least when the unlensed power spectra are sufficiently smooth functions.[136]

4.4 Summary

In this Chapter, we have presented a systematic, complete full-sky harmonic-space formalism for weak-lensing distortions to the CMB temperature and polarization fields, as well as their auto-/cross-correlation spectra. We have taken advantage of the TAM formalism (also see Ref.[137] for a recent review). Our results are valid for arbitrary linear metric perturbations in an FLRW background in general gauges. We have provided the complete weak-lensing effects on polarization, which includes a subtle correlated sum between the deflection of line of sight and the gravitational Faraday rotation of polarization. Together with the perturbation to photon redshift, we have exhausted the list of effects that metric perturbations can impact on the propagation of photons.

Our new contribution to this extensively studied subject is the analysis of gravitational Faraday rotation of the photon polarization, which may be induced by linear vector or tensor metric perturbations in the Universe. It should be emphasized that polarization rotation should not be regarded as a separate effect from gravitational deflection. As we have explained, the common origin from spacetime perturbations implies a crucial, full correlation between the two, so that their impact on the CMB statistics has to be included simultaneously. Despite non-detection of the curl deflection in the CMB data, curl deflection estimators have been developed to monitor systematics[118] and to put limits on primordial vector perturbation and tensor perturbation.[138,139] To correctly extract information from polarization maps, the cross correlation between curl deflection and rotation has to be consistently accounted for. To this end, previous estimators constructed to measure direction-

dependent rotation $\psi(\hat{\mathbf{n}})$, although motivated by considerations of non-gravitational new physics, can be directly borrowed.[140–143] Therefore, curl deflection estimators have to be supplemented with rotation estimators under a joint analysis to put robust limits on exotic gravitational lenses beyond the scalar gravitational potential generated by inhomogeneous matter distribution.

Our analysis has been carried out at linear order in metric perturbations. This corresponds to the Born approximation, for which the total lensing effect is the sum of independent lens sheets along the line of sight. This means that the multiple-deflection couplings between different lens sheets have been neglected. Under this approximation, we have shown that the polarization rotation ψ is identical to the image rotation, which is given by the anti-symmetric part of the deflection Jacobian, the image distortion tensor. For realistic lensing by foreground large-scale structure, however, lens-lens coupling in the quasi-weak or strong lensing regime may be important and a second-order computation is in need.[144–148] It would be interesting to extend the analysis to second order to uncover possible general relation between the polarization rotation ψ and the image distortion. We would like to leave that for future work.[149]

Chapter 5

Imprints From Gravitational Waves in Large-Scale Structure

Complementary to the fluctuations in the CMB, large-scale structure of the Universe provides another powerful avenue to study cosmological perturbations and their initial condition. Its observational scope covers the large-scale distribution of dark matter and baryons, as well as various types of astrophysical objects that have formed through gravitational collapse. It spans the epoch of structure formation during matter domination, from the completion of cosmic recombination $z < 1100$, till the very recent dark-energy-dominated Universe $z \lesssim 1$. It probes cosmological length scales that range from mildly perturbed scales $100 - 10^4$ Mpc to highly nonlinear scales $\sim$ kpc $-$ Mpc. The central task underlying the study of the large-scale structure is to make inference about the initial perturbations generated via the inflation mechanism, through measurement of the matter

distribution or tracer clustering at recent redshifts.

Thanks to rapid development in instrumental technology in astronomy, there has been a major push in the observation of the large-scale structure, and hence a growing ambition to deepen our understanding about the Universe through detailed studies. Current or planned large-scale structure surveys aim at larger angular and deeper redshift coverage, better imaging resolution, more diversity of tracers with different clustering properties and formation physics, as well as a multi-component characterization of the sky that spans decades in the wavelength. Despite that in many observations the implication for early Universe is expected to be obscured by astrophysical complexity, the three dimensional nature of large-scale structure observations guarantees a great wealth of cosmological information to be uncovered.

An almost scale-invariant background of cosmic-scale gravitational waves is a key prediction of single-field slow-roll inflation.[30–33] Active experimental efforts to detect signatures in CMB B-mode polarization, generated by the part of the gravitational-wave spectrum that entered the horizon after recombination, has been underway.[53,55] For a slow-roll phase with an expansion e-folds $N \sim 50 - 60$, it may also be possible to directly detect solar-scale relic gravitational waves using ground-based or space laser interferometry in the future.[150–153]

On the other hand, there have been proposals on detecting signatures of primordial gravitational waves in large-scale structure surveys, based on its lensing distortion to the apparent galaxy clustering,[154] cosmic shear from gravitational lensing,[155,156] or distortion to the apparent 21cm brightness fluctuation.[87,157] However, those have been generally

regarded not promising in terms of detection prospects, for the following two reasons. First, gravitational waves are decoupled from dark matter clustering in linear theory of large-scale structure formation, because in the linearized theory, gravitational waves are tensor perturbations to the spacetime metric, while dark matter density perturbation is sourced by scalar gravitational potentials. Furthermore, there is no sustained growth in the wave amplitude throughout the epoch of structure formation during matter domination, when no significant anisotropic stress in the cosmic fluid is present. On the contrary, linear gravitational wave amplitude decays as the Universe continues to expand, as long as the wavelength is on subhorizon scale. For these reasons, no impact on the physical clustering process was considered before. Rather, attention was mainly given to the kinetic effect of gravitational waves on photon propagation, which is not significant, both due to the smallness of subhorizon wave amplitude, and due to the propagating nature of the wave.[158]

The impact of gravitational waves on the physical clustering of dark matter had remained largely unexplored, except for a previous work by Masui and Pen[159] in which the idea of a fossil quadrupolar anisotropy was put forth for the first time. The effect being described here is small, as in the language of perturbation theory it arises from a coupling between the scalar density perturbation and the tensor metric perturbation at second order. Still, there is no reason to take it less seriously than the gravitational lensing effect, which is also of second order in perturbation.

From the observational point of view, a clean target effect would most likely arise from large-scale gravitational waves affecting the growth of short-scale density modes. We will be focusing on this kinematic limit hereafter. There are two aspects of the physics. First,

the inflation mechanism sets up in the initial condition an apparent modulation of short-scale scalar perturbations with large-scale tensor modes, which is fixed by the requirement that no physical modulation is existant when power spectrum is evaluated at fixed physical scales rather than coordinate scales. Second, after perturbation modes enter the horizon and start to evolve, short-scale density waves locally experience a tidal force from large-scale gravity waves, which breaks the isotropy of the collapse dynamics. The net result is a locally observable anisotropic clustering that may be measured in upcoming large-scale structure surveys. Based on our work Ref.,[59] this Chapter will be devoted to detailed explanation and analysis on this subject.

We will be considering the following FLRW metric perturbed by scalar perturbation (in the Poisson gauge[129,160]) and tensor perturbation (i.e. gravitational waves)

$$\begin{aligned} ds^2 &= -\left[1+2\Phi(\mathbf{x},t)\right]dt^2 + 2a(t)w_i(\mathbf{x},t)dx^i dt \\ &\quad + a^2(t)\left[1-2\Psi(\mathbf{x},t)\right]\left[\delta_{ij}+\gamma_{ij}(\mathbf{x},t)\right]dx^i dx^j, \end{aligned} \tag{5.1}$$

where the tensor metric perturbation γ_{ij} is traceless and divergence-free, and for consistency we have included a vector perturbation w_i (satisfying $\partial^i w_i = 0$) induced only at second order. We will use uppercase $\mathbf{K}$ for tensor wave vector and lowercase $\mathbf{k}_1, \mathbf{k}_2, \cdots$ for scalar modes, and always assume a hierarchy in scales $K \ll k_1, k_2, \cdots$.

We will first introduce in Section 5.1 the tensor-scalar-scalar consistency relation for standard single-field inflation, and discuss its physical significance. Then in Section 5.2, we introduce the concept of the gravitational-wave "fossil" imprint using a toy model of

mass lattice in an expanding Universe. Next we present a rigorous calculation of the fossil effect during the epoch of large-scale structure formation in Section 5.3, using perturbation theory. Based on that result, Section 5.4 will be on the detection of fossil imprints in galaxy clustering surveys, with emphasis on the importance of projection effects when defining a gauge invariant galaxy observable. Finally, concluding comments will be given in Section 5.5.

5.1 Tensor-scalar-scalar consistency relation

Single-field slow-roll inflation (SFSR) produces both scalar adiabatic perturbations and tensor perturbations. The second-order mode coupling can be most conveniently quantified by a tensor-scalar-scalar bispectrum in the form of Eq. (3.3). In the post-inflationary epoch when all cosomogical perturbation modes of interest are well beyond the Hubble scale, Maldacena first derived that *in the squeezed limit* $K \ll k = |(\mathbf{k}_2 - \mathbf{k}_1)/2|$ and $k \approx k_1 \approx k_2$, the bispectrum is simply related to the scalar spectral slope[20,161]

$$\langle \Phi_{\rm ini}(\mathbf{k}_1)\Phi_{\rm ini}(\mathbf{k}_2)\gamma_{{\rm ini},\alpha}(\mathbf{K})\rangle = (2\pi)^3\delta_D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{K})\,\frac{1}{2}\frac{d\ln P_\Phi}{d\ln k}\epsilon_\alpha^{ij}(\mathbf{K})\hat{k}_{1i}\hat{k}_{2j}P_\gamma(K)P_\Phi(k), \quad (5.2)$$

where $d\ln P_\phi/d\ln k = n_s - 4$, and the label $_{\rm ini}$ signifies primordial perturbation amplitude on superhorizon scales. Here $\epsilon_\alpha^{ij}(\mathbf{K})$ is the polarization tensor for a plane wave of wave vector $\mathbf{K}$ and polarization state α, and $P_\gamma(K)$ and $P_\Phi(k)$ are the linear tensor and scalar power spectra, respectively. Note that the model-dependent correction due to a finite ratio K/k is typically $\mathcal{O}[(K/k)^2]$ and hence negligible. This squeezed bispectrum is interpreted as a

spatial modulation of scalar power spectrum of short wavelengths, calculated locally within a patch (around $\mathbf{x}$) smaller than the tensor wavelength, by the local tensor perturbation value $\gamma_{\mathrm{ini},ij}(\mathbf{x})$,

$$P_\Phi(k)|_{\gamma_{\mathrm{ini}}(\mathbf{x})} = P_\Phi(k)\left[1 - \frac{1}{2}\frac{d\ln P_\Phi}{d\ln k}\gamma_{\mathrm{ini},ij}(\mathbf{x})\hat{k}^i\hat{k}^j\right], \tag{5.3}$$

Does this result imply that one could infer the existence of arbitrarily long-wavelength gravitational waves by a local measurement of power anisotropy in the scalar sector? If yes, it would be contradicting the famous *Equivalence Principle* in general relativity, which says that local measurements do not reveal the existence of gradual variation in spacetime geometry.[59,159,162]

The resolution of the paradox is that it is only meaningful to ask whether or not the scalar (or matter) power spectrum evaluated at *fixed physical scale* is being modulated. The wave number k in Eq. (5.3), conjugate to a spatial scale $\lambda \sim 2\pi/k$, is a *coordinate separation*. According to the metric Eq. (5.1), the proper spatial separation is given by $d\tilde{x}^i = (\delta^i{}_j + \frac{1}{2}(\gamma_{\mathrm{ini}})^i{}_j)dx^j$, corresponding to a proper wave vector

$$\tilde{k}^i = \left(\delta^i{}_j - \frac{1}{2}(\gamma_{\mathrm{ini}})^i{}_j\right)k^j. \tag{5.4}$$

A linearized calculation then clearly shows that when measured at a fixed proper scale k,

$$\begin{aligned} P_\Phi(\tilde{k})\Big|_{\gamma_{\rm ini}(\mathbf{x})} &= P_\Phi\left(\left(\delta^i{}_j + \frac{1}{2}(\gamma_{\rm ini})^i{}_j(\mathbf{x})\right)k^j\right)\Big|_{\gamma_{\rm ini}(\mathbf{x})} \\ &= P_\Phi(k)\left[1 - \frac{1}{2}\frac{d\ln P_\Phi}{d\ln k}\gamma_{{\rm ini},ij}(\mathbf{x})\hat{k}^i\hat{k}^j + \frac{1}{2}\frac{d\ln P_\Phi}{d\ln k}\gamma_{{\rm ini},ij}(\mathbf{x})\hat{k}^i\hat{k}^j\right] \\ &= P_\Phi(k), \end{aligned} \tag{5.5}$$

the power is *not* modulated by the tensor mode. We see that in a general relativistic framework, although the tensor-scalar-scalar consistency relation Eq. (5.2) is a correct mathematical result, its interpretation needs caution. It will be exactly the physical scale, rather than the coordinate scale (which does depend on one's choice of coordinate system), that specific physical processes in the late-time universe are dependent on.

Therefore, the consistency relation for the squeezed bispectrum Eq. (5.2) sets up the correct initial condition for tensor-scalar coupling at second order, in such a way that there is no *physical* correlation between long-wavelength tensor modes and short-wavelength scale modes [1]. This results from the physical picture that when the long-wavelength modes had been excited from the de Sitter vacuum and had long been stretched to scales much larger than the horizon size, the short-wavelength adiabatic modes were excited as if they were in a homogeneous inflationary universe.

[1] The same argument and interpretation carry through for the squeezed scalar-scalar-scalar consistency relation.[163]

5.2 Fossil imprint of gravitational waves: a comsic lattice analogy

If the formation of galaxies depend on the scalar power spectrum at given proper scale [2], does Eq. (5.5) imply that there will be no physical modulation of galaxy clustering that can be observed locally? The answer is no for a measurement done at later times when structure formation has taken place, because gravitational waves are dynamical after horizon re-entry! A simple argument was first given in Ref.[159] We demonstrate this effect using the toy model as described below.

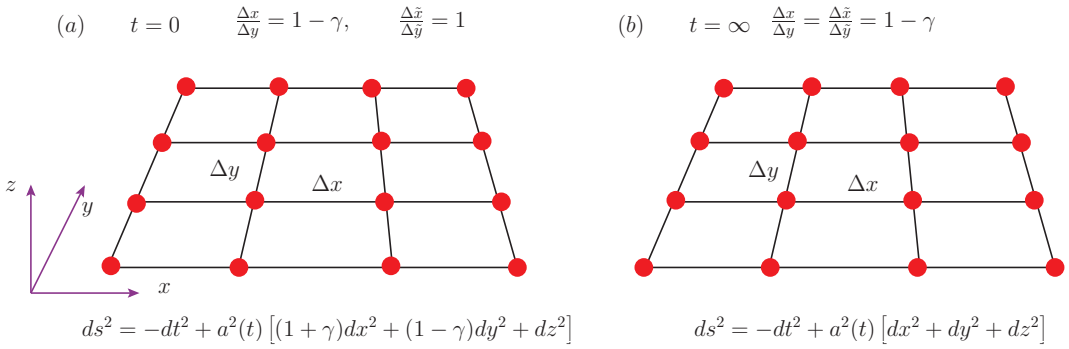


Figure 5.1: We use a toy model to demonstrate the fossil imprint of inflationary gravitational waves on the distribution of matter in the Universe.

Imagine a homogeneously expanding Universe populated by a bunch of massive test particles, which form a regular, three-dimensional lattice. This is sketched in Figure 5.1, where we show a *comoving* patch of the two-dimensional lattice living on the $x-y$ plane, at two different times. Let Δx and Δy be the *coordinate* lattice spacing along the x direction and the y direction, respectively. Similarly, let $\Delta\tilde{x}$ and $\Delta\tilde{y}$ be the *proper* lattice spacing along

[2] In cosmology we usually refer to given comoving scale. However, it merely differs from the proper scale by the homogeneous scale factor $a(t)$. The conversion between the two therefore is not crucial to our argument.

the two perpendicular directions. Imagine that at early times (left panel of Figure 5.1), the Universe is perturbed by a gravitational wave, which propagates along the z direction and has wavelength much longer than the horizon scale. Since the wave amplitude (assumed to be small $\gamma \ll 1$ so that we can linearize with respect to it) freezes out in this regime, the metric can be written as

$$\text{Early time}: \qquad ds^2 = -dt^2 + a^2(t)\left[(1+\gamma)dx^2 + (1-\gamma)dy^2 + dz^2\right], \tag{5.6}$$

i.e. a stretching-squashing mode along the two transverse directions. If the primordial dynamics during the inflationary epoch is assumed to be isotropic, an analogue of inflationary tensor-scalar-scalar consistency relation Eq. (5.2) would dictate that a local observer would expect a square lattice to begin with, if measured in proper length. This requires that $\Delta\tilde{x}/\Delta\tilde{y} = 1$ and hence $\Delta x/\Delta y = 1-\gamma$, i.e. unequal coordinate lattice spacing. Now we examine what the toy Universe looks like at sufficiently late times, when the Universe is in a decelerating phase and the gravitational wave will have its wavelength much shorter than the Hubble scale. The Universe will have become homogeneous FLRW,

$$\text{Late time}: \qquad ds^2 = -dt^2 + a^2(t)\left[dx^2 + dy^2 + dz^2\right], \tag{5.7}$$

because the wave amplitude will have decayed significantly. However, solving geodesic motion to first order in γ leads to the conclusion that all test masses will have fixed comoving coordinate $\mathbf{x}$, which implies constant Δx and Δy for the cosmic lattice. What will have changed is the mapping between the coordinate separation and the proper separation, due

to the change in the spatial metric. The proper lattice spacing ratio at late times is

$$\left.\frac{\Delta\tilde{x}}{\Delta\tilde{y}}\right|_{t=\infty} = \left.\frac{\Delta x}{\Delta y}\right|_{t=\infty} = 1-\gamma, \tag{5.8}$$

indicating that a local observer will see a rectangular lattice at later times, a breakdown of isotropy.

From the perspective of a local observer, who always measures proper length scales, the fossil effect results from the action of a time-dependent tidal force induced locally by the dynamical gravitational wave. Although the tidal acceleration is suppressed by the size of the lattice with respect to the wavelength of the tensor mode, it causes a non-negligible accumulative displacement on the test mass.

The aforementioned imprint of gravitational waves has a few nice features. First, the effect is *not* suppressed by the ratio between the short scale of the problem and the long scale related to gravitational waves. As a matter of fact, the effect is independent of the short scale, as long as the short scale is much shorter compared to the long scale. We have not specified how large the lattice spacing scale has to be when compared to the gravitational wavelength; it can be arbitrarily small. Second, the level of anistropy induced in the lattice is proportional to the initial (superhorizon) wave amplitude, rather than the amplitude at any later time. For this reason, the effect is referred to as a "fossil" effect, in the sense that it is permanently imprinted in the matter distribution, even after gravitational waves have decayed away (see also the memory effect of gravitational waves[164–167]). In this regard, accurate measurement of the late-time matter distribution will allow for a reconstruction

of the initial condition for inferring the existence of any primordial gravitational-wave perturbation.

The toy model we have constructed here is of course an oversimplification of the physics. There is no lattice in the real Universe, but merely a random distribution of mass. However, specific correlation scales in the matter field (even the primodial adibatic perturbation may have a scale invariant spectrum, the resultant matter power spectrum is not), resulting from a specific structure growth history in given cosmology, can play the role of a cosmic stick that measures statistical isotropy along different directions. Next, we invoke the general relativistic perturbation theory to study the realistic case of a matter-dominated Universe with initial perturbations set up according to the prediction of standard single-field inflation.

5.3 Fossil imprint of gravitational waves: perturbative approach

We consider an Einstein-de Sitter (EdS) universe with flat geometry ($\Omega_m = 1$), which is anticipated to be a good approximation when considering tensor perturbation modes that enter the horizon after the onset of matter domination. Neglecting baryons, non-relativistic matter is then assumed to be described by a pressure-free fluid of cold dark matter alone. At very low redshifts, a dark energy component emerges and starts to dominate the expansion history. We ignore this realistic complication in order to be able to derive simple analytical results. In any case, adopting this simplification does not change the qualitative aspects of our conclusion.

We start with the metric Eq. (5.1) and perturbatively solve the effect of long-wavelength tensor perturbation γ_{ij} on short-scale matter clustering. The latter is quantified by two scalar gravitational potentials Φ and Ψ, the matter density fluctuation $\delta = \rho_m/\bar{\rho}_m - 1$, and a peculiar velocity field v^i for the cold dark matter fluid. Since the fossil effect is second order in perturbation, in principle we have to consistently derive and solve gravitational and fluid equations to that order. To this end, we decompose perturbations into a first-order solution (labeled with $^{(1)}$ and whose solution is well known from the linear perturbation theory) and a second-order solution (labeled with $^{(2)}$) due to tensor-scalar coupling. For instance, we decompose the potential perturbation $\Phi = \Phi^{(1)} + \Phi^{(2)}$, and we do the same for Ψ, δ, γ_{ij}, and so on. For the peculiar velocity $v_i = v_i^{(1)} + v_i^{(2)} + v_{R,i}$, the first-order part is a purely gradient flow $v_i^{(1)}$ of cold gravitational collapse, but the second-order part contains both a gradient flow $v^{(2)}$ and a vortical flow $\partial^i v_{R,i} = 0$. Full second-order calculation in cosmological perturbation theory is a sophisticated subject. However, we may ignore terms quadratic in linear tensor perturbation $\gamma_{ij}^{(1)}$, or terms quadratic in linear scalar/matter perturbations, which do not contribute to tensor-scalar coupling at second order. Since we are also not interested in seeking the second-order correction $\gamma_{ij}^{(2)}$, a negligible backreaction from the scalar/matter perturbations on the gravitational waves,we shall use γ_{ij} throughout to represent the linear tensor mode. Following these rules, a perturbative expansion for the curvature tensor as well as the cold dark matter energy-stress tensor are summarized in Appendix C.

5.3.1 Linear evolution

The knowledge of linear evolution is needed to derive second-order corrections. The linear solutions in EdS are well known. At linear order, scalar/matter perturbations are decoupled from gravitational waves. The matter peculiar velocity is curl-free $v_i^{(1)} = \partial_i v^{(1)}$. In the absence of linear anisotropic stress, the two scalar gravitational potentials coincide $\Psi^{(1)} = \Phi^{(1)}$. Assuming adiabatic initial condition from standard inflation, the potential is constant throughout matter domination $\Phi^{(1)}(\mathbf{x}, t) = \Phi^{(1)}(\mathbf{x}, 0) = \Phi_{\rm ini}(t)$. The growth of cold dark matter density and velocity are given by

$$\delta^{(1)} = -2\mathcal{T}_\delta(k)\Phi_{\rm ini}, \quad v^{(1)} = -\frac{2}{3aH}\Phi_{\rm ini}, \tag{5.9}$$

where the linear growth factor for matter is given by $\mathcal{T}_\delta(k) = 1 + k^2/(3a^2H^2)$. On the other hand, a linear gravitational wave with wave number K obeys the following wave equation

$$\ddot{\gamma}_{ij} + 3H\dot{\gamma}_{ij} + a^{-2}K^2\gamma_{ij} = 0. \tag{5.10}$$

The solution is a linear extrapolation from the initial amplitude in the superhorizon regime $\gamma_{ij} = \mathcal{T}_\gamma(K)\gamma_{{\rm ini},ij}$, where in EdS we simply have $\mathcal{T}_\gamma(K, \tau) = 3j_1(K\tau)/(K\tau)$ at conformal time τ.

5.3.2 Second order tensor-scalar coupling

As detailed in Appendix C.2, we subtract the linear part from the perturbed Einstein equations Eqs. (C.12)–(C.14) and fluid equations Eqs. (C.15)–(C.16) and keep the terms of second order in perturbation. This leads to the same second-order differential equations for second-order variables themselves as their first-order counterparts, except that in addition they are sourced by inhomogeneous source terms that are quadratic in linear perturbations. According to our perturbation scheme, those have to be the product of linear tensor mode and linear scalar/matter modes.

It is most convenient to first derive an equation for the second-order potential $\Psi^{(2)}$,

$$\ddot{\Psi}^{(2)} + 4H\dot{\Psi}^{(2)} = \frac{\partial^{-2}}{a^2}\left[\left(\partial^2\gamma_{ij}\right)\left(\partial^i\partial^j\Phi^{(1)}\right)\right] + H\dot{\Gamma}^{(2)}, \tag{5.11}$$

with $\Gamma^{(2)} = \Psi^{(2)} - \Phi^{(2)}$ given by

$$\partial^2\Gamma^{(2)} = 3\partial^{-2}\left[\left(\partial^2\gamma^{ij}\right)\left(\partial_i\partial_j\Phi^{(1)}\right)\right]. \tag{5.12}$$

Assuming much longer gravitational wavelength compared to the scale of cold dark matter perturbations $K \ll k$, the inverse Laplacian can be approximated by

$$\partial^{-2}\left[\left(\partial^2\gamma^{ij}\right)\left(\partial_i\partial_j\Phi^{(1)}\right)\right] \approx \left(\partial^2\gamma^{ij}\right)\left(\partial^{-2}\partial_i\partial_j\Phi^{(1)}\right), \tag{5.13}$$

with a negligible error suppressed by $\mathcal{O}[(K/k)^2]$. With the scale hierarchy, the second source

term on the right hand side of Eq. (5.11) is in fact also a factor $\mathcal{O}[(K/k)^2]$ smaller than the first term, and therefore may be dropped. Eq. (5.11) can be solved by the method of Green's function. Ignoring the source terms first, the homogeneous equation has a constant solution $f_1(t) = 1$ and a decaying solution $f_2(t) = [a(t)H(t)]^5$. The retarded Green's function then reads

$$\begin{aligned} G_{\rm ret}\left(t-t'\right) &= \frac{f_1(t)f_2(t') - f_2(t)f_1(t')}{\dot{f}_1(t')f_2(t') - \dot{f}_2(t')f_1(t')}\Theta\left(t-t'\right) \\ &= \frac{2}{5H(t')}\left[1-\left(\frac{a(t)H(t)}{a(t')H(t')}\right)^5\right]\Theta\left(t-t'\right), \end{aligned} \tag{5.14}$$

The solution for $\Psi^{(2)}$ is

$$\begin{aligned} \Psi^{(2)}(t) &= \int_0^t dt' \left[a^{-2}(t')\left(\partial^2\gamma^{ij}(t')\right)\left(\partial^{-2}\partial_i\partial_j\Phi^{(1)}(t')\right)\right] G_{\rm ret}\left(t-t'\right) \\ &= -\mathcal{S}(K)\,\gamma^{ij}_{\rm ini}\left(\frac{\partial_i\partial_j}{\partial^2}\Phi_{\rm ini}\right), \end{aligned} \tag{5.15}$$

where we define a function

$$\mathcal{S}(K) = \int_0^{K\tau} d(K\tau')\frac{K\tau'}{5}\mathcal{T}_\gamma(K\tau')\left[1-\left(\frac{K\tau'}{K\tau}\right)^5\right]. \tag{5.16}$$

As shown in Figure 5.2, the function $\mathcal{S}(K)$ starts off from zero when the gravitational wave mode is superhorizon ($K\tau \to 0$), and asymptotes to 3/5 when the wave has damped out sufficiently ($K\tau \gg 1$), representing a residual "fossil" effect.

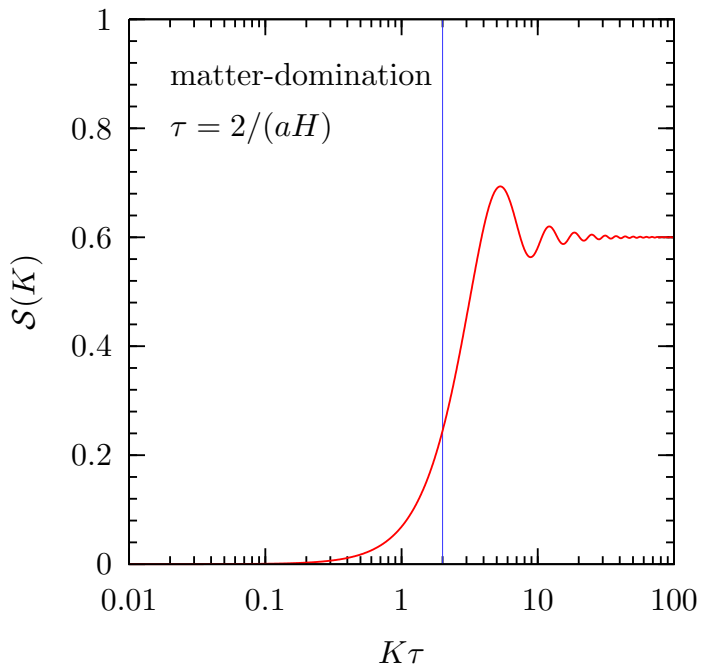


Figure 5.2: The function $\mathcal{S}(K)$ during matter domination, which depends on the product $K\tau$. The moment of horizon entry $K\tau = 2$ has been shown with the vertical line.

Using Eq. (C.27), the density field corrected at second order is found to be

$$\delta = -2\mathcal{T}_\delta(k)\left(1 - \frac{1}{2}\frac{d\ln\mathcal{T}_\delta(k)}{d\ln k}\mathcal{T}_\gamma(K)\gamma^{ij}_{\rm ini}\hat{k}_i\hat{k}_j\right)\Phi_{\rm ini} + 2\mathcal{T}_\delta(k)\gamma^{ij}_{\rm ini}\hat{k}_i\hat{k}_j\Phi_{\rm ini}\mathcal{S}(K). \quad (5.17)$$

5.3.3 Second order tensor-scalar coupling: Lagrangian perspective

In the previous section, we have used the Eulerian approach to solve how gravitational waves affect the evolution of dark matter density field and leave a fossil imprint. Here we take the Lagrangian point of view to derive the same result. For simplicity, we shall focus on dark matter clustering on subhorizon scales $k \gg aH$.

The Lagrangian point of view is particularly intuitive for cold dark matter clustering in

the weakly perturbed region where shell crossing of dark matter paticle flow have not yet occured. In this case, a bunch of dark matter particles found at some position $\mathbf{x}$ at time t can be uniquely traced back to some initial position $\mathbf{q}$, i.e. the *Lagrangian position*, before dark matter particles have started to cluster under gravity. Therefore, the evolution of the dark matter distribution amounts to solving the displacement field $\mathbf{s}(\mathbf{q}, t)$, defined through

$$\mathbf{x}(\mathbf{q}, t) \quad = \quad \mathbf{q} + \mathbf{s}(\mathbf{q}, t). \tag{5.18}$$

At initial time $t \to 0$, it may be assumed that the density perturbations are negligibly small, so that dark matter particles are uniformly distributed in Lagrangian space. The density contrast at later times is then given by the volume conversion between the physical space $d^3\mathbf{x}$ and the Lagrangian space $d^3\mathbf{q}$,

$$\delta \quad = \quad \det\left[\frac{\partial x^i}{\partial q^j}\right]^{-1} - 1 \approx -\partial \cdot s. \tag{5.19}$$

For the approximation mark, we perturbatively expand the volume Jacobian to linear order in perturbation. Note that to account for tensor-scalar coupling at second order, we do not need an expansion for the Jacobian to second order. Because tensor perturbation conserves the volume at linear order, any second order correction to the volume distortion only arises from the second-order matter clustering itself, unrelated to the tensor perturbation. According to this perturbation scheme, when spatial derivatives act on first order variables (e.g. the displacement $\mathbf{s}$), it is not necessary to distinguish between the Lagrangian position $\mathbf{q}$ and the Eulerian position $\mathbf{x}$. Since dark matter particles are free falling, the covariant

displacement $s^\mu = (t, \mathbf{s})$ traces out a geodesic

$$\frac{d^2 s^\mu}{d\lambda^2} = -\Gamma^\mu_{\alpha\beta} \frac{ds^\alpha}{d\lambda} \frac{ds^\beta}{d\lambda}, \tag{5.20}$$

parametrized by the proper time λ. It suffices to identify λ with the coordinate time t in our case, because γ_{ij} does not perturb g_{00} and thus induces no dilation in proper time. The Christoffel symbols are found in Eq. (C.7), except that we shall set $\Psi = \Phi$ in the Newtonian limit. To close the system, we further provide the Newtonian Poisson equation (valid on subhorizon scales) that determines how the gravitational potential Φ is sourced by the distribution of dark matter particles in physical space

$$\frac{1}{a^2} \left(\delta^{ij} - \gamma^{ij}\right) \partial_i \partial_j \Phi = 4\pi G \rho_m \delta = 4\pi G \rho_m \left(-\partial \cdot s\right). \tag{5.21}$$

It can be seen that gravitational waves have the geometrical effect that the "physical" Laplacian entering the Poisson equation should be $(\delta^{ij} - \gamma^{ij})\partial_i \partial_j$ (see the time-time Einstein equation Eq. (C.12)), due to the "proper distance" argument we have put forth in Section 5.1.

The displacement $\mathbf{s}$ can be solved perturbatively. We split $s^i = s^{(1)i} + s^{(2)i}$, where $s^{(1)i}$ is the displacement in the absence of γ_{ij}, and $s^{(2)i}$ corrects the Lagrangian trajectory for the effect of γ_{ij}. As can be easily derived, the first-order trajectory is determined by the familiar Newton's formula for acceleration in a gravitational potential

$$\ddot{s}^{(1)i} = -\frac{1}{a^2} \partial^i \Phi^{(1)}. \tag{5.22}$$

We may take the divergence, and then combine with the Poisson equation to eliminate $\Phi^{(1)}$. By doing so, we derive a second-order equation for the divergence of the displacement

$$\frac{d}{dt}\left[a^3\frac{d}{dt}\left(\frac{\partial\cdot s^{(1)}}{a}\right)\right]+a^3H\frac{d}{dt}\left(\frac{\partial\cdot s^{(1)}}{a}\right) = 0, \tag{5.23}$$

which admits a growing solution during matter domination [3]

$$s^{(1)i}=-\frac{2}{3a^2H^2}\partial^i\Phi^{(1)}=-\frac{2}{3a^2H^2}\partial^i\Phi_{\rm ini}. \tag{5.24}$$

Now we proceed to solve for $s^{(2)i}$. We start with the spatial geodesic equation up to second order

$$\frac{d^2s^i}{dt^2} = -\Gamma^i_{00}-2\Gamma^i_{j0}\frac{ds^j}{dt}-\Gamma^i_{jk}\frac{ds^j}{dt}\frac{ds^k}{dt}, \tag{5.25}$$

and then extract the part at second order in perturbation

$$\ddot{s}^{(2)i} = -\frac{1}{a^2}\partial^i\Phi^{(2)}-\frac{1}{a}\left(\dot{w}^i+Hw^i\right)+\frac{1}{a^2}\gamma^{ij}\partial_j\Phi^{(1)}-2H\dot{s}^{(2)i}-\dot{\gamma}^i_j\dot{s}^{(1)j}. \tag{5.26}$$

Again, the divergence of this equation reads

$$\frac{d}{dt}\left[a^2\frac{d}{dt}\left(\partial\cdot s^{(2)}\right)\right] = -\partial^2\Phi^{(2)}+\gamma^{ij}\partial_i\partial_j\Phi^{(1)}-\frac{2}{3H}\dot{\gamma}^{ij}\partial_i\partial_j\Phi^{(1)}, \tag{5.27}$$

where we have used divergence-free conditions $\partial\cdot w=0$ and $\partial^i\gamma_{ij}=0$. The second-order

[3] Note that the linear displacement generated by gravitational clustering is curl-free.

part of the Poisson equation Eq. (5.21) reads

$$\frac{1}{a^2}\partial^2\Phi^{(2)} - \frac{1}{a^2}\gamma^{ij}\partial_i\partial_j\Phi^{(1)} \;=\; -4\pi G\rho_m \partial\cdot s^{(2)}. \tag{5.28}$$

Eq. (5.27) can be combined with Eq. (5.28) for elimination of the gravitational potential. We obtain in terms of the conformal time $d\tau = dt/a$

$$\frac{d^2}{d\tau^2}\left(\partial\cdot s^{(2)}\right) + \frac{2}{\tau}\frac{d}{d\tau}\left(\partial\cdot s^{(2)}\right) - \frac{6}{\tau^2}\left(\partial\cdot s^{(2)}\right) \;=\; \frac{2}{3H}\dot{\gamma}^{ij}\partial_i\partial_j\Phi_{\rm ini}. \tag{5.29}$$

The general structure for the solution to Eq. (5.29) may be written

$$\partial\cdot s^{(2)} \;=\; \left(\partial\cdot s^{(2)}\right)_{\rm homo} + \left(\partial\cdot s^{(2)}\right)_{\rm spec}, \tag{5.30}$$

where $\left(\partial\cdot s^{(2)}\right)_{\rm homo}$ is a solution to the homogeneous part, and $\left(\partial\cdot s^{(2)}\right)_{\rm spec}$ accounts for the nonzero source term on the right hand side. The former can be found by considering the superhorizon limit $K\to 0$, for which γ_{ij} reduces to a frozen, constant amplitude. Then we have

$$\left(\partial\cdot s^{(2)}\right)_{\rm homo} \;=\; -\frac{2}{3a^2H^2}\gamma_p^{ij}\partial_i\partial_j\Phi_{\rm ini}, \tag{5.31}$$

in line with the inflation consistency relation Eq. (5.39). The inhomogeneous solution $\left(\partial\cdot s^{(2)}\right)_{\rm spec}$ can be solved using the method of Green's function. For matter domination, the homogeneous solution has a growing mode $g_1(\tau)=\tau^2$ and a decaying mode $g_2(\tau)=\tau^{-3}$,

and the *retarded* Green's function reads

$$G_{\rm ret}\left(\tau-\tau'\right) = \frac{g_1(\tau)g_2(\tau')-g_2(\tau)g_1(\tau')}{g_1'(\tau')g_2(\tau')-g_2'(\tau')g_1(\tau')}\Theta\left(\tau-\tau'\right) = \frac{\tau^2}{5\tau'}\left[1-\left(\frac{\tau'}{\tau}\right)^5\right]\Theta\left(\tau-\tau'\right), \quad (5.32)$$

where a prime represents derivative with respect to τ. Then we find the solution to be

$$\begin{aligned}\left(\partial\cdot s^{(2)}\right)_{\rm spec} &= \int_0^\tau d\tau'\left(\frac{2}{3H(\tau')}\dot\gamma^{ij}(\tau')\partial_i\partial_j\Phi_{\rm ini}\right)G_{\rm ret}\left(\tau-\tau'\right)\\ &= \frac{2}{3a^2H^2}\mathcal{S}_N(K)\gamma^{ij}_{\rm ini}\partial_i\partial_j\Phi_{\rm ini},\end{aligned} \quad (5.33)$$

where we have defined a new function

$$\mathcal{S}_N(K) \equiv \int_0^{K\tau} d(K\tau')\,\frac{2}{5}\frac{\partial\mathcal{T}_\gamma(K\tau')}{\partial(K\tau')}\left[1-\left(\frac{K\tau'}{K\tau}\right)^5\right], \quad (5.34)$$

which is related to $\mathcal{S}(K)$ through

$$\mathcal{S}_N(K)+\frac{1}{2}\frac{d\ln\mathcal{T}_\delta}{d\ln k}\left(1-\mathcal{T}_\gamma(K)\right)=\mathcal{S}(K), \quad (5.35)$$

with $d\ln T_\delta/d\ln k = 2$ for matter domination. Note that $\mathcal{S}_N(K)$ also starts off from zero at early times $K\tau \ll 1$ and then asymptotes to $-2/5$ at $K\tau\to\infty$. The final result for the density contrast is found to be

$$\begin{aligned}\delta &= -\partial\cdot s^{(1)}-\partial\cdot s^{(2)}\\ &= -2\mathcal{T}_\delta(k)\Phi_{\rm ini}\left[1-\frac{1}{2}\frac{d\ln\mathcal{T}_\delta}{d\ln k}\gamma_{{\rm ini},ij}\hat k^i\hat k^j-\mathcal{S}_N(K)\,\gamma_{{\rm ini},ij}\,\hat k^i\hat k^j\right].\end{aligned} \quad (5.36)$$

This result agrees with Eq. (5.17) in the subhorizon limit $k \gg aH$, for which we may take $T_\delta(k) = k^2/(3a^2H^2)$.

The Lagrangian computation of the fossil imprint is an improvement of the toy model of Section 5.2 to the realistic problem of cold dark matter clustering.

5.3.4 Consistency relation for density field

As a special limit of the central result Eq. (5.17), we consider the case of infinite tensor wavelength $K \to 0$, or equivalently an early moment well before horizon entry. Then $\mathcal{T}_\gamma(K) = 1$ and $\mathcal{S}(K) = 0$. The answer for density field is reduced to

$$\delta = -2\mathcal{T}_\delta(k)\left(1 - \frac{1}{2}\frac{d\ln\mathcal{T}_\delta(k)}{d\ln k}\gamma^{ij}_{\rm ini}\hat{k}_i\hat{k}_j\right)\Phi_{\rm ini}. \tag{5.37}$$

This leads to a local density two-point correlation function in the presence of the very long-wavelength tensor mode

$$\begin{aligned}\langle\delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\rangle_{\gamma_{\rm ini}} &= \int d^3k\, e^{i\mathbf{k}\cdot\mathbf{x}} 4\mathcal{T}_\delta^2(k)\left(1 - \frac{1}{2}\frac{d\ln\mathcal{T}_\delta^2(k)}{d\ln k}\gamma^{ij}_{\rm ini}\hat{k}_i\hat{k}_j\right) P_\Phi(\mathbf{k})|_{\gamma_{\rm ini}} \\ &= \int d^3k\, e^{i\mathbf{k}\cdot\mathbf{x}} 4\mathcal{T}_\delta^2(k)\left(1 - \frac{1}{2}\frac{d\ln\mathcal{T}_\delta^2(k)}{d\ln k}\gamma^{ij}_{\rm ini}\hat{k}_i\hat{k}_j\right)\left(1 - \frac{1}{2}\frac{d\ln P_\Phi}{d\ln k}\gamma^{ij}_{\rm ini}\hat{k}_i\hat{k}_j\right) P_\Phi(k) \\ &= \int d^3k\, e^{i\mathbf{k}\cdot\mathbf{x}}\left(1 - \frac{1}{2}\frac{d\ln P_\delta}{d\ln k}\gamma^{ij}_{\rm ini}\hat{k}_i\hat{k}_j\right) P_\delta(k),\end{aligned} \tag{5.38}$$

according to $P_\delta(k) = 4\mathcal{T}_\delta^2(k)P_\Phi(k)$. A local matter power spectrum modulated by the very long-wavelength tensor mode can be directly read off

$$P_\delta(k)|_{\gamma_{\rm ini}(\mathbf{x})} = P_\delta(k)\left[1 - \frac{1}{2}\frac{d\ln P_\delta}{d\ln k}\gamma_{{\rm ini},ij}(\mathbf{x})\hat{k}^i\hat{k}^j\right], \tag{5.39}$$

which is analogous of Eq. (5.3). Following the same reasoning as in Eq. (5.5), the interpretation of this result is that a local observer measures precisely *no* correlation between the density fluctuation power on fixed proper scale and the long-wavelength gravitational wave.

For finite gravitational wavelength K, local observer measures a physical modulation of local matter power spectrum by the gravitational wave, originating from the second term $\propto \mathcal{S}(K)$ in Eq. (5.17). This is an effect from the tidal force, an observable consequence of the oscillating wave that cannot be attributed to an improper choice of spatial coordinates. This is also not a transient effect, because that term asymptotes to a finite value for $K\tau \to \infty$ even though in that limit the tensor amplitude will die off $\mathcal{T}_\gamma \to 0$.

5.4 Fossil imprints in galaxies

Even if galaxies perfectly trace the invisible dark matter density perturbation, i.e. assuming no galaxy clustering bias, Eq. (5.17) does not represent the observed galaxy distribution. In fact, in a inhomogeneous Universe, Eq. (5.17) describes the density field on a slice of constant time in given gauge, namely a given choice of spacetime coordinates. The freedom to choose different spacetime coordinates implies that the definition of constant-

time slice is gauge dependent, and hence is the answer for a perturbation field such as Eq. (5.17). To eliminate this *gauge ambiguity*, a gauge-invariant galaxy observable needs to be defined.

5.4.1 Observed galaxy location

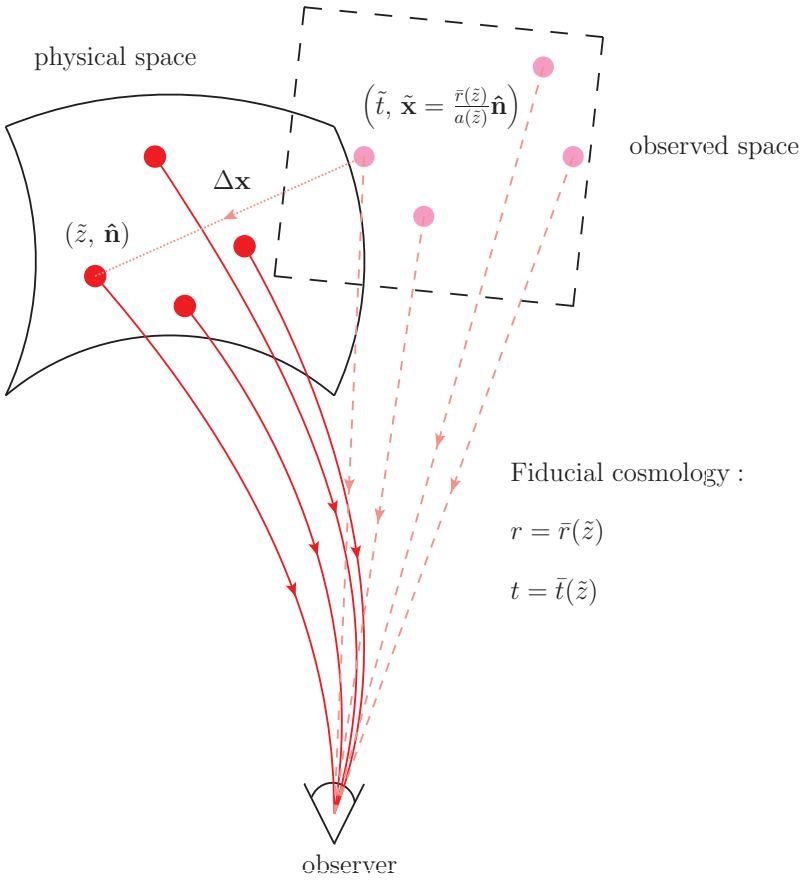


Figure 5.3: The apparent distribution of galaxies in observed space (dashed) differs from the intrinsic distribution in physical space (solid) due to photon projection effects. The distortion is described by a displacement vector $\Delta\mathbf{x} = \tilde{\mathbf{x}} - \mathbf{x}$.

Consider a galaxy, located in a perturbed Universe at comoving position $\mathbf{x}$ and cosmic time t in some globally defined coordinate system. In realistic galaxy redshift surveys,

however, only two quantities about this galaxy may be measured — the *observed* (spectroscopy or photometric) redshift $\tilde{z}$, and its *apparent* direction $\hat{\mathbf{n}}$ on the sky. The mapping between $(\tilde{z}, \hat{\mathbf{n}})$ and its physical coordinates $(t, \mathbf{x})$ is unknown. The observer may only make an inference about the location of this galaxy under the assumption of a specific homogeneous cosmology. This means that there is a background radial-distance–redshift relation $r = \bar{r}(z)$, a background time–redshift relation $t = \bar{t}(z)$, and a background scale-factor–redshift relation $a = a(z)$. The inferred galaxy position and source time are then given by

$$\tilde{\mathbf{x}} = \frac{\bar{r}(\tilde{z})}{a(\tilde{z})}\hat{\mathbf{n}}, \qquad \tilde{t} = \bar{t}(\tilde{z}). \tag{5.40}$$

This procedure defines a mapping between the distribution of galaxies in physical spacetime and the distribution of galaxies in observed coordinates (Figure 5.3). The difference between the physical location $(t, \mathbf{x})$ and the apparent location $(\tilde{t}, \tilde{\mathbf{x}})$, quantified by a (small) displacement $\Delta\mathbf{x} = \mathbf{x} - \tilde{\mathbf{x}}$ and $\Delta t = t - \tilde{t}$, is resultant from distortion to photon propagation by the inhomogeneous spacetime, i.e. kinetic or gravitational redshift/blueshift of the photon energy and gravitational deflection, as has been discussed in Section 4.1. Since it is only the observed distribution rather than the physical distribution that is directly measurable, the observed galaxy number density is a valid gauge invariant observable.

Projection effects can induce first-order galaxy perturbation in two ways. First, a given physical location is mapped to a location in observed space corresponding to a different cosmic time. Since galaxy number density constrast is defined with respect to some global

mean at the inferred cosmic time, the observed density constrast is changed. Second, a given volume in physical spacetime, including both the spatial comoving volume and the redshift interval, is changed when mapped to the observed coordinates. This volume distortion effect, quantified by a transformation Jacobian, also alters the observed galaxy number density. For systematic analysis of the projection effects in galaxy surveys, we refer readers to Ref.[56,168–173] However, what is relevant to the fossil effect here is the additional modulation of galaxy clustering through a distortion to photon propagation by the gravitational wave, which is in the form of a tensor-scalar mode coupling at second order. In this case, instead of modifying the local galaxy number density, projection effects from tensor modes modify the local galaxy two-point correlation function.[174] It is therefore worth clarifying that what needs to be computed for our purpose is the projection effect from long-wavelength gravitational waves, rather than that from short-wavelength scalar/matter perturbations. This means that when computing the projection effects, we consider the following spacetime

$$ds^2 = -dt^2 + a^2(t)\left(\delta_{ij} + \gamma_{ij}\right) dx^i dx^j, \tag{5.41}$$

perturbed only by (the long-wavelength) γ_{ij}.

The spatial and temporal displacements $\Delta\mathbf{x}$ and Δt have been computed up to first order in perturbation, through an integral along the *unperturbed* line of sight back to the

source redshift. The results are given by[121,154]

$$\Delta t = \frac{1}{2H}\int_0^{\chi_e} d\chi\,\frac{\partial\gamma_{\parallel}}{\partial\tau}, \tag{5.42}$$

$$\Delta x_{\parallel} = -\frac{1}{2}\int_0^{\chi_e} d\chi\,\gamma_{\parallel} - \frac{1}{2aH}\int_0^{\chi_e} d\chi\,\frac{\partial\gamma_{\parallel}}{\partial\tau}, \tag{5.43}$$

$$\Delta x^i_{\perp} = \frac{\chi_e}{2}\left(\gamma^{ij}_o\hat{n}_j - \gamma_{o,\parallel}\hat{n}^i\right) + \Pi^{ij}\int_0^{\chi_e} d\chi\left(\frac{\chi_e-\chi}{2}\partial_j\gamma_{\parallel} - \frac{\chi_e}{\chi}\hat{n}^k\gamma_{jk}\right), \tag{5.44}$$

where the integration variable χ is the comoving radial distance and χ_e is the comoving radial distance to the source galaxy. We have decomposed Δx^i into a line-of-sight component and a transverse component $\Delta x^i = \hat{n}^i\Delta_{\parallel} + \Delta x^i_{\perp}$ with $\Pi^i{}_j\Delta x^j_{\perp} = 0$. We have also introduced the notation $\gamma_{\parallel} = \gamma_{ij}\hat{n}^i\hat{n}^j$. The time delay Δt and line-of-sight displacement $\Delta x_{\parallel}$ is a result of the kinetic and the gravitational redshift/blueshift, while the transverse displacement $\Delta x^i_{\perp}$ is the well known gravitational deflection. Note that for infinite tensor wavelength $K \to 0$, Δt vanishes and $\Delta x^i = -\gamma^{ij}_{\rm ini}\,x_j/2$.

5.4.2 Observed galaxy two-point correlation

We now derive the observed galaxy two-point correlation function. Throughout, we will assume an unbiased galaxy population on large-scales $\delta_g = \delta_m$. Since a general relativistic definition for galaxy bias is by itself a subtle subject worthy of separate study,[163,168,175,176] we adopt this simplification, though unrealistic, to circumvent the subtlety of a consistent theory of bias in general relativity. We will also consider an idealistic galaxy survey in which we do not worry about practical selection effects such as magnification bias, evolution bias, and so on. This then means that galaxy number is conserved when the physical galaxy

distribution is being mapped to the observed one. We therefore have

$$a^3\left(\tilde{t}\right)\tilde{n}_g\left(\tilde{\mathbf{x}},\tilde{t}\right)d^3\tilde{x} = a^3(t)\left[\det\left(\delta_{ij}+\gamma_{ij}\right)\right]^{1/2} n_g(\mathbf{x},t)d^3x, \tag{5.45}$$

Of course, linear tensor perturbation does not physically perturb the volume element $\det(\delta_{ij}+\gamma_{ij}) = 1$. Furthermore, in the absence of an evolution bias, comoving galaxy number density is constant [4], and therefore $a^3\left(\tilde{t}\right)\bar{n}_g\left(\tilde{t}\right) = a^3(t)\bar{n}_g(t)$. Using this fact, combining with the definition of galaxy perturbation in both physical coordinates and observed coordinates,

$$n_g\left(\mathbf{x},t\right) = \bar{n}_g(t)\left[1+\delta_g\left(\mathbf{x},t\right)\right], \tag{5.46}$$

$$\tilde{n}_g\left(\tilde{\mathbf{x}},\tilde{t}\right) = \bar{n}_g(\tilde{t})\left[1+\tilde{\delta}_g\left(\tilde{\mathbf{x}},\tilde{t}\right)\right], \tag{5.47}$$

we obtain

$$\tilde{\delta}_g - \delta_g = \left(\Delta_t\partial_t + \Delta x^i\partial_i\right)\delta_g + \partial_i\Delta x^i + \delta_g\,\partial_i\Delta x^i. \tag{5.48}$$

We may ignore the second term, which merely accounts for the additional linear clustering proportional to the tensor perturbation and thus has nothing to do with tensor-scalar mode coupling. The time derivative in the first term may also be neglected if we focus on subhorizon matter clustering $k \gg aH$.

The observed galaxy power spectrum on small scales is the Fourier transform of the

[4] This is also an oversimplification in reality, because galaxies or other matter tracers may form or perish, and therefore their comoving number density may evolve with time. This is conventionally quantified by an evolution bias $b_e = d\ln(a^3\bar{n}_g)/d\ln a$.

two point correlation function

$$\tilde{P}_g(\mathbf{k}) = \int d^3\delta\mathbf{x}\, e^{-i\mathbf{k}\cdot\delta\mathbf{x}} \left\langle \tilde{\delta}_g(\mathbf{x}_1)\, \tilde{\delta}_g(\mathbf{x}_2) \right\rangle_{\gamma(\mathbf{x})}. \tag{5.49}$$

Here we are correlating two points $\mathbf{x}_1 = \mathbf{x} - \delta\mathbf{x}/2$ and $\mathbf{x}_2 = \mathbf{x} + \delta\mathbf{x}/2$, separated by a short distance from the midpoint $\mathbf{x} = (\mathbf{x}_1 + \mathbf{x}_2)/2$ compared to the tensor wavelength. Therefore, while galaxy number density fluctuates significantly from $\mathbf{x}_1$ to $\mathbf{x}_2$, the tensor perturbation γ_{ij} may be approximated as homogeneous around $\mathbf{x}$. We now want to insert Eq. (5.48) into each of the two density fields on the right hand side of Eq. (5.49). We must account for the fact that the two density fields are evaluated at two distinct points $\mathbf{x}_1$ and $\mathbf{x}_2$. Using

$$\frac{\partial}{\partial\mathbf{x}_1} = \frac{1}{2}\frac{\partial}{\partial\mathbf{x}} - \frac{\partial}{\partial\delta\mathbf{x}}, \qquad \frac{\partial}{\partial\mathbf{x}_2} = \frac{1}{2}\frac{\partial}{\partial\mathbf{x}} + \frac{\partial}{\partial\delta\mathbf{x}}, \tag{5.50}$$

to convert $\mathbf{x}_1$ and $\mathbf{x}_2$ to the midpoint location $\mathbf{x}$ and separation $\delta\mathbf{x}$, we have a leading-order expansion for the $\Delta x^i \partial_i \delta_g$ term,

$$\begin{aligned}
\Delta x^i(\mathbf{x}_1)\frac{\partial}{\partial x_1^i}\delta(\mathbf{x}_1) &= \left(\Delta x^i - \frac{1}{2}\delta x^j \partial_j \Delta x^i\right)\left(\frac{1}{2}\frac{\partial}{\partial x^i} - \frac{\partial}{\partial \delta x^i}\right)\delta(\mathbf{x}_1), \\
\Delta x^i(\mathbf{x}_2)\frac{\partial}{\partial x_2^i}\delta(\mathbf{x}_2) &= \left(\Delta x^i + \frac{1}{2}\delta x^j \partial_j \Delta x^i\right)\left(\frac{1}{2}\frac{\partial}{\partial x^i} + \frac{\partial}{\partial \delta x^i}\right)\delta(\mathbf{x}_2),
\end{aligned} \tag{5.51}$$

where the slowly varying functions Δx^i and $\partial_j \Delta x^i$ are to be evaluated at the midpoint $\mathbf{x}$ and are hence treated as independent of $\delta\mathbf{x}$. Then we have the following expressions for the

Fourier transform of the derivative terms

$$\begin{aligned}\Delta x^i(\mathbf{x}_1)\frac{\partial}{\partial x_1^i}\delta(\mathbf{x}_1) &= \int d^3\mathbf{k}e^{i\mathbf{k}\cdot(-\delta\mathbf{x}/2)}\left(\Delta x^i + i\partial_j\Delta x^i\frac{\partial}{\partial k_j}\right)\left(\frac{1}{2}\frac{\partial}{\partial x^i}+\frac{i}{2}k_i\right)\delta(\mathbf{k}),\\ \Delta x^i(\mathbf{x}_2)\frac{\partial}{\partial x_2^i}\delta(\mathbf{x}_2) &= \int d^3\mathbf{k}e^{i\mathbf{k}\cdot(\delta\mathbf{x}/2)}\left(\Delta x^i + i\partial_j\Delta x^i\frac{\partial}{\partial k_j}\right)\left(\frac{1}{2}\frac{\partial}{\partial x^i}+\frac{i}{2}k_i\right)\delta(\mathbf{k}).\end{aligned} \quad (5.52)$$

The term $\delta_g\,\partial_i\Delta x^i$ is easy to handle because it suffices to evaluate $\partial_i\Delta x^i$ at the midpoint $\mathbf{x}$. When inserting Eq. (5.48) and Eq. (5.52) into Eq. (5.49), we encounter ensemble average around $\mathbf{x}$ in the presence of tensor perturbation $\langle\cdots\rangle_{\gamma(\mathbf{x})}$. However, if such a term is multiplied by coefficients of first order in γ_{ij}, we may insert the global linear galaxy power spectrum $\langle\delta_g(\mathbf{k})\delta_g(\mathbf{k}')\rangle = (2\pi)^3\delta(\mathbf{k}+\mathbf{k}')P_g(k)$. This is independent of the direction of $\mathbf{k}$, and is also independent of the location of the local patch $\mathbf{x}$. These results allow us to derive the observed short-scale galaxy power spectrum

$$\tilde{P}_g(\mathbf{k})\Big|_{\gamma(\mathbf{x})} = P_\delta(\mathbf{k})|_{\gamma(\mathbf{x})} - \left(\partial_j\Delta x_i\right)(\mathbf{x})\,\hat{k}^i\hat{k}^j\frac{d\ln P_\delta(k)}{d\ln k}P_\delta(k) + \left(\partial_i\Delta x^i\right)(\mathbf{x})\,P_\delta(k). \quad (5.53)$$

Using the perturbative result we have obtained in Eq. (5.17) for the physical matter clustering with second-order tensor-scalar coupling, we finally obtain

$$\begin{aligned}\tilde{P}_g(\mathbf{k})\Big|_{\gamma(\mathbf{x})} = P_g(k)\Bigg[1 &- \frac{d\ln P_\Phi(k)}{d\ln k}\hat{k}^i\hat{k}^j\left(\partial_j\Delta x_i+\frac{1}{2}\gamma_{\mathrm{ini,ij}}\right)\\ &-2\frac{d\ln\mathcal{T}_\delta(k)}{d\ln k}\hat{k}^i\hat{k}^j\left(\partial_j\Delta x_i+\frac{1}{2}\gamma_{ij}\right) - 2\mathcal{S}(K)\gamma^{ij}_{\mathrm{ini}}\hat{k}_i\hat{k}_j\Bigg].\end{aligned} \quad (5.54)$$

The significance of this result is that in the presence of large-scale gravitational waves, the observed galaxy power spectrum on small scales receives a direction-dependent correction

upon the isotropic, globally averaged linear galaxy power spectrum. The correction is in the form of a linear power modulation, proportional to the amplitude of long-wavelength gravitational waves. The directional dependence of this piece is quadrupular, and it has alignment with the shearing directions of the background gravitational wave. This modulation of short-scale power spectrum by large-scale modes is a clear indication of an measurable nonzero tensor-galaxy-galaxy bispectrum $\langle\delta_g(\mathbf{k}_1)\delta_g(\mathbf{k}_2)\gamma(\mathbf{K})\rangle$ in the squeezed limit $K \ll k_1, k_2$.

In the limit of infinite tensor wavelength $K \to 0$, the correction in Eq. (5.54) vanishes. This is expected, because a constant γ_{ij} can be eliminated by an anisotropic rescaling of the spatial coordinates and cannot have any physical consequence. Applying this to the observable universe, gravitational waves whose wavelengths are still longer than the present Hubble scale $K \ll H_0$ leave no observable imprint whatsoever. However, it is worth noting that the undetectability of superhorizon gravitational waves crucially relies on the inflationary tensor-scalar-scalar consistency relation Eq. (5.3). It is conceivable that new physical mechanisms during inflation might have generated genuine correlation between a long-wavelength tensor mode γ_{ij} and short-wavelength adiabatic modes.[60,61,177] Then the SFSR consistency relation Eq. (5.3) can be violated, and the fossil imprint in Eq. (5.54) would not necessarily vanish for very long tensor wavelength.

5.4.3 Position-dependent galaxy power quadrupole

The anisotropic part of the observed galaxy power spectrum, Eq. (5.54), has a quadrupolar dependence $\sim \gamma^{ij}_{\rm ini}\hat{k}_i\hat{k}_j$ on the direction of small-scale clustering $\hat{\mathbf{k}}$, induced by the physical

coupling between matter perturbation and gravitational waves, and some additional, misaligned quadrupolar dependence $\sim (\partial_j \Delta x_i)\hat{k}^i\hat{k}^j$ resultant from photon propagation. The former part aligns with the metric strain tensor γ_{ij}, while the latter part also depends on the line-of-sight direction $\hat{n}$ through $\partial_j \Delta x_i$. This angular dependence, modulated in a position dependent way by the gravitational wave configuration, is a characteristic departure from a universe without any large-scale relic gravitational wave, for which the small-scale matter power spectrum is expected to be isotropic.

In order to quantify the quadrupolar anisotropy, we define the following (normalized) quadrupole moments,

$$\mathcal{Q}_{2m}(\mathbf{x}) \equiv \frac{\int d^2\hat{\mathbf{k}}\, \tilde{P}_g(\mathbf{k};\mathbf{x}) Y^*_{(2m)}(\hat{\mathbf{k}})}{\int d^2\hat{\mathbf{k}}\, \tilde{P}_g(\mathbf{k};\mathbf{x}_c) Y^*_{(00)}(\hat{\mathbf{k}})}, \tag{5.55}$$

for $m = \pm 2, \pm 1, 0$. These five quadrupole moments will transform by a rotation matrix if the quantization axis for $Y_{(lm)}(\hat{\mathbf{k}})$ is changed, and the choice of quantization axis is arbitrary. This can be done for every local patch smaller than the tensor wavelength and therefore the quadrupole moments are *position-dependent*, i.e. they vary with the central location $\mathbf{x}$. These harmonic moments may be computed through

$$\mathcal{Q}_{2m}(\mathbf{x}) = \int d^2\hat{\mathbf{k}}\, \mathcal{Q}_{ij}(\mathbf{x}) \left(\hat{k}^i\hat{k}^j - \frac{1}{3}\delta^{ij}\right) Y^*_{(2m)}(\hat{\mathbf{k}}). \tag{5.56}$$

According to Eq. (5.54), the position-dependent power quadrupole tensor, being symmetric

and trace-free by construction, is given by

$$\begin{aligned}\mathcal{Q}_{ij}(\mathbf{x}) &= -\frac{1}{2}\frac{d\ln P_\delta}{d\ln k}\left(1-\mathcal{T}_\gamma\right)\gamma_{\mathrm{ini},ij}(\mathbf{x}) + 2\mathcal{S}_N(K)\gamma_{\mathrm{ini},ij}(\mathbf{x}) \\ &\quad -\frac{d\ln P_\delta}{d\ln k}\left[\frac{1}{2}\gamma_{ij}(\mathbf{x}) + \partial_{(i}\Delta x_{j)}(\mathbf{x}) - \frac{1}{3}\delta_{ij}\ \left(\partial\cdot\Delta x\right)(\mathbf{x})\right]. \qquad (5.57)\end{aligned}$$

The first term describes the fossil imprint that is exactly captured by the toy model of Section 5.2 ("Lagrangian fossil effect"). The second term may be interpreted as the influence of gravitational waves on the displacement from gravitational infall (see Section 5.3.3), which is not modeled in Section 5.2. The final term is the gauge invariant photon projection contribution with the proper inclusion of "metric shear".[155,156] It can be directly verified that any contribution to the physical observable $\mathcal{Q}_{ij}(\mathbf{x})$ from gravitational waves with superhorizon wavelength vanishes.

Due to the stochastic nature of the inflationary gravitational-wave background, the observed galaxy power quadrupole moments are random. We may still predict the *mean square* quadrupole

$$\overline{\mathcal{Q}^2} \equiv \left\langle \sum_{m=-2}^{2} |\mathcal{Q}_{2m}|^2 \right\rangle, \qquad (5.58)$$

where we average over the gravitational-wave ensemble. This is an invariant observable under a rotation of coordinate system

$$\overline{\mathcal{Q}^2} = (8\pi/15)\left\langle \mathcal{Q}_{ij}\mathcal{Q}^{ij*}\right\rangle. \qquad (5.59)$$

If we expand the Gaussian gravitational-wave background using the TAM formalism (Eq. (4.79)) with a power spectrum *per linear polarization state* Eq. (4.81), we find the mean square galaxy power quadrupole at given source redshift z to be

$$\overline{\mathcal{Q}^2}(z) = \frac{4}{15\pi}\int_{K_{\rm min}}^{K_{\rm max}} K^2 dK P_\gamma(K) \sum_{J=2}^{\infty}(2J+1)\mathcal{Q}_J^2(K,z), \tag{5.60}$$

after summing over independent TAM eigenmodes labeled by the wave number K and the total angular momentum J, M (summation over M can be explicitly done due to statistical isotropy). For a self-contained presentation, we present explicit expressions for the kernels $Q_J^2(K,z)$ in Appendix D. In practice, the summation over J can be truncated at $J \sim K\chi$, where χ is the comoving radial distance to the source redshift z. This is because a TAM wave with a large total-angular-momentum has no support at small radii $\chi \lesssim J/K$. Physical cutoffs for the range of gravitational wavenumber K are also required. Since we have demonstrated that superhorizon wavelengths leave no physical imprint, the long-wavelength end is physically truncated at the present horizon scale $K_{\rm min} \sim H_0$. On the other hand, the short-wavelength end should be subject to the condition $K_{\rm max} \ll k$ if our analysis assuming a scale hierarchy is to apply.

5.4.4 Numerical results

We now conduct a numerical study for the standard inflation scenario in which the gravitational-wave background has an almost scale invariant power spectrum $P_\gamma(K) = 2\pi^2\Delta_\gamma^2/K^3$. We then assume that the late-time cosmology is described by the flat ΛCDM

concordance model with the *WMAP*+BAO+H_0 best-fit cosmological parameters.[96]

The perturbative results of Section 5.3 have been derived assuming a matter-dominated universe. This is an oversimplification even for the discussion of large-scale structure formation. First of all, the concordance cosmology predicts that the Universe is dominated by radiation before $z \sim 4000$, during which the growth of cold dark matter perturbation is refrained. This results in a power break in the matter power spectrum around the horizon scale at matter-radiation equality $k = k_{\rm eq}$. Indeed, galaxy surveys probe scales that enter the horizon before that point, the growth of which is not described by the gravitational clustering dynamics in the EdS universe.

During radiation domination, the diffusion length of photons in the coupled baryon-photon plasma sets the scale of *Silk damping*. On larger scales, radiation perturbation dominates the gravitational potential, and therefore one needs to solve the coupled system of photon-baryon acoustic waves and cold dark matter overdensity to derive the fossil effects. When the former dominates, its oscillatory nature results in no coherent fossil imprint. Below the Silk damping scale, however, the photon-baryon acoustic waves quickly damp out, and therefore the gravitational potential becomes dominated by the cold dark matter. In this case, it suffices to solve the growth of dark matter perturbation in a *homogeneous* baryon-photon plasma. To this end, our analysis in Section 5.3 can be generalized to the case of a background expansion dominated by radiation.[178] Here we will present a simpler computation by ignoring *all* fossil imprints that are possibly generated prior to the epoch of matter domination. This means that the matter power spectrum at matter-radiaton

equality will be treated as the "initial" power spectrum, which amounts to setting

$$\mathcal{T}_\delta(k) \simeq \begin{cases} k^2/(3a^2H^2), & k < k_{\rm eq}, \\ k_{\rm eq}^2/(3a^2H^2), & k > k_{\rm eq}. \end{cases} \tag{5.61}$$

On the other hand, we shall focus on large-scale gravitational waves that enter the horizon after matter-radiation equality. The solution for the linear gravitational-wave amplitude in the EdS universe then continues to apply. To ensure a scale hierarchy between gravitational waves and matter perturbations, we will only include $K < k_{\rm eq}$.

At recent redshifts $z < 2$, the Universe is observed to be Dark-Energy-dominated. This accelerates the cosmic expansion and therefore affects the dynamics of both matter perturbation and gravitational waves. However, the correction is expected to be marginal. We can take the analytical results derived for the EdS universe, express temporal evolution everywhere as a function of the scale factor $a(\tau)$, and then use the appropriate $a(z)$ or $a(\tau)$ relations for ΛCDM cosmology. This approximation turns out to be fairly precise in the presence of an emerging Dark Energy component.[160]

Figure 5.4 provides a demonstration of how gravitational waves on different scales contribute to the mean square observed galaxy power quadrupole $\overline{\mathcal{Q}^2}$. Naively, Maldacena's tensor-scalar-scalar consistency relation for single-field inflation would imply a uniform distribution per logarithmic interval of K for a scale-invariant primordial tensor power spectrum. This even extends to tensor wavelengths that are far beyond the present-day horizon scale. As we have explained before, however, Maldacena's consistency relation does not

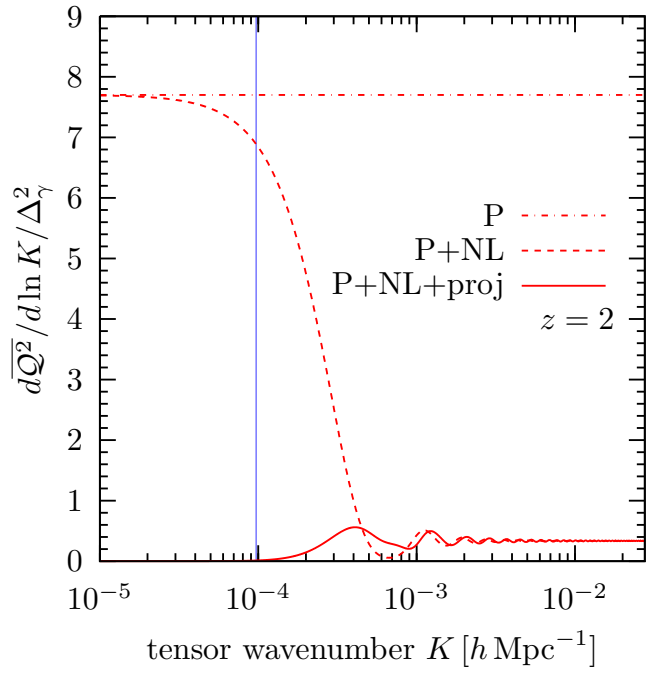


Figure 5.4: Mean square galaxy power quadrupole $\overline{\mathcal{Q}^2}$ from *per logarithmic interval* of gravitational wavenumber K for a source redshift $z = 2$. The result is normalized to Δ_γ^2. We compare three cases: (1) if only the inflation tensor-scalar-scalar consistency relation is included (dash-dotted); (2) if the nonlinear tensor-scalar mode coupling is accounted for in the dark matter clustering (dashed); (3) gauge invariant observable with the full photon projection effects included (solid). The vertical line represents the present-day horizon scale H_0.

correspond to a gauge-invariant observable. The observable galaxy power quadrupole has to include nonlinear tensor-scalar mode coupling after horizon entry, and relativistic projection effects on photon projection. We see that superhorizon gravitational waves $K < H_0$ do not leave observable fossil imprint, in spite of Maldacena's consistency relation. We also see that the contribution asymptotes to a constant for subhorizon wavelengths $K \gg H_0$, indicating that the galaxy quadrupole is dynamically built up and is a permanent imprint in the galaxy distribution.

In Figure 5.4, we sum over a wide range of scales in the gravitational-wave power

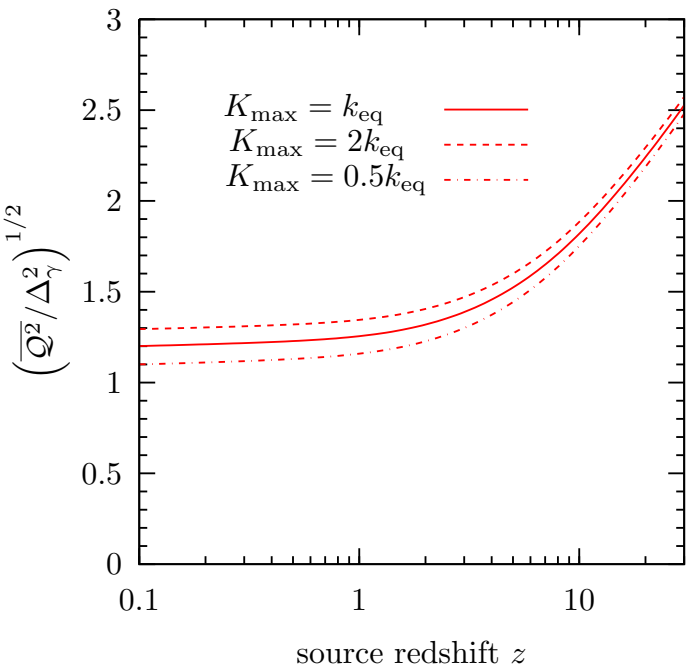


Figure 5.5: Root mean galaxy power quadrupole as a function of source redshift z. The result is normalized to Δ_γ^2. Marginal change in the results is found when $K_{\rm max}$ varies.

spectrum and find the total root mean galaxy power quadrupole for given source redshift z. For a scale-invariant primordial tensor power spectrum, the result is not sensitive to the specific wave number cutoff adopted. It can be seen that for all redshifts relevant for current or forthcoming large-scale structure surveys, the root mean galaxy quadrupole is expected to be of the order of primordial tensor root mean amplitude $(\Delta_\gamma^2)^{1/2}$. Given the current constraint on the inflationary tensor-to-scalar ratio $r \lesssim 0.1$, the prospect of detecting fossil imprints from primordial gravitatonal waves is severely limited by cosmic variance. A promising detectability may be achieved by high-redshift cosmic 21cm surveys in the future.

5.5 Summary

In this Chapter, we have studied the imprints of large-scale relic gravitational waves in the large-scale structure of the Universe. We have shown that although the single-field inflation tensor-scalar-scalar bispectrum implies no physical correlation between long-wavelength gravity waves and short-wavelength matter perturbations, and hence the absence of modulation of the initial matter statistics, a physical power quadrupole in galaxy clustering arises in a position-dependent manner, once the gravitational wave becomes dynamical after horizon entry. This "fossil" signature, at the order of the primordial gravitational-wave amplitude, is permanently imprinted in the distribution of galaxies, despite that gravitational waves decay due to cosmic expansion.

In standard cosmology, position-dependent galaxy power quadrupole is expected to result from second-order nonlinear gravitational collapse, and from gravitational lensing by foreground large-scale structure. Line-of-sight galaxy power quadrupole is known to arise from redshift space distortions.[179] All those effects are at a level probably much greater than that of gravitational waves. Is the gravity-wave fossil signature completely indistinguishable from other second-order effects? Not quite! According to Eq. (D.1), parity-odd wave eigenmodes $\alpha = TB$ induce a position dependent power quadrupole that varies across the sky according to the two parity-odd tensor spherical harmonics $Y^{VB}_{(JM)ij}(\hat{\mathbf{n}})$ and $Y^{TB}_{(JM)ij}(\hat{\mathbf{n}})$, while other effects from large-scale scalar perturbations do *not* produce those parity-odd patterns. One may take advantage of this property to devise optimal estimators to look for unique signatures of the gravitational waves.

The fossil signature from standard single-field inflation are likely to be too small to be detected in cosmic-variance-limited surveys. However, if the single-field inflation consistency relation for tensor-scalar-scalar bispectrum does not hold in the early Universe, a much larger galaxy power quadrupole may be expected to correlate with relic gravitational waves, even for wavelengths longer than the size of the visible Universe! Tight constraints on those alternative mechanism of gravitational-wave generation can be derived from a non-detection of galaxy power quadrupole.

Furthermore, fossil imprint of gravitational waves may have a consequence on the shape alignment of galaxies. Galaxy weak lensing surveys have revealed a large-scale correlation between galaxy shear (aligned ellipticity of the observed galaxy shape) and large-scale density distribution around those galaxies .[180–184] Since cosmic shear induced by gravitational lensing by foreground galaxies along is unable to explain the observed correlation, the result is interpreted as an *intrinsic* alignment of galaxy shape with the surrounding large-scale structure. Although it is not entirely transparent how the large-scale surrounding environment precisely impacts the process of galaxy formation and thus induces the shape alignment, it is generally believed that this intrinsic alignment results from the tidal force generated by the large-scale matter distribution.[185–194] Therefore, it is conceivable that tidal field generated by long-wavelength gravitational waves during the epoch of galaxy formation may also lead to galaxy shape alignment.[155,178]

To distinguish between the contribution from gravitational waves and from gravitational potentials, the spin-two shear map is conventionally decomposed into the parity-even E-modes and the parity-odd B-modes, analogous to what has been done for the CMB po-

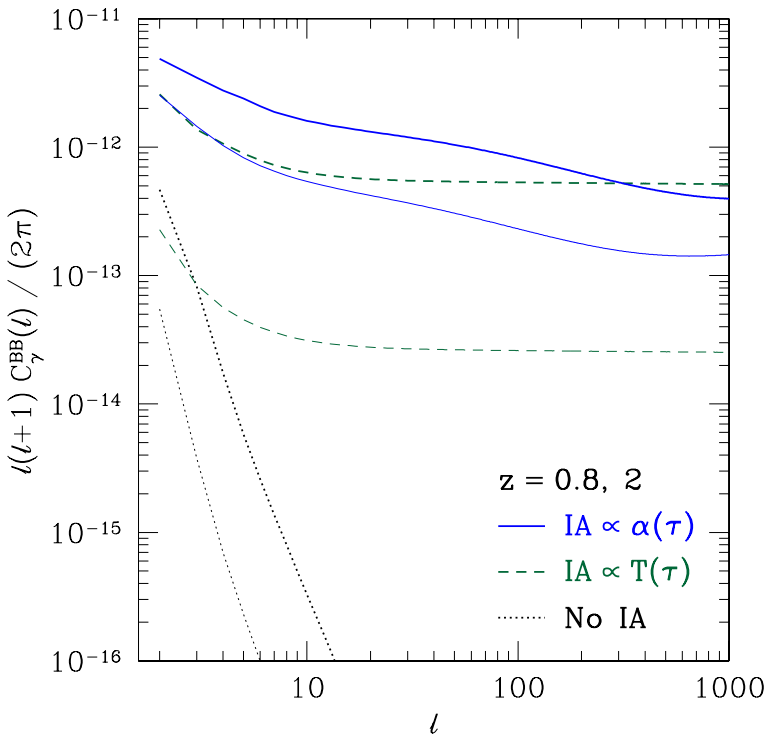


Figure 5.6: Parity-odd cosmic shear induced by tidal influence on galaxy formation from gravitational waves. The fossil alignment scenario (blue solid) predicts stronger alignment than the instantaneous alignment scenario (green dashed) over a wide range of angular scales. Figure is taken from Ref.[178]

larization. Since the parity-odd signal is not expected to be generated by the gravitational potential at linear order, it serves as a smoking gun for gravitational waves. This can open up a new possibility to search for possible signatures of gravitational waves in the large-scale structure. As an example, Figure 5.6 shows the angular power spectrum for the B-mode intrinsic alignment for a tensor-to-scalar ratio $r = 0.1$, at two different source redshifts, with an interesting signal level on large angular scales. The statistical significance, mainly limited by the random shape noise, may be greatly reduced with high galaxy number density in forthcoming galaxy surveys such as *Euclid*[195] and *LSST*.[196] Finally, intrinsic shear from gravitational waves is also expected to correlate with large-scale CMB B-mode polarizations

from the reionization epoch.[197] It would be interesting to develop more accurate and realistic models for the galaxy intrinsic alignment with large-scale gravitational waves, either from first principles of the underlying physics, or relying on fast computational simulations.

Chapter 6

Separate Universe Formalism

Accurately modeling the large-scale structure of the Universe is in general considered a greater theoretical challenge than modeling the cosmic microwave background radiation. While the latter can be described by linearized equations of motion to very good precision, the former inevitably evolves into a regime of nonlinear dynamics due to gravitational instability, on relatively short length scales $\lesssim 100$ Mpc in the recent Universe. Given the plentiful of cosmological information contained in the cosmic large-scale structure about the physical state of the very early Universe, which we have no means to observe directly, even incremental progress in understanding the nonlinear aspect of large-scale structure formation will be very rewarding.

The extremely large spatial and temporal scales involved in the physics of large-scale structure create a special realm where one may test the general relativistic theory of gravity still in the weak field regime. Indeed, the Newtonian theory of gravity is expected to be

incapable of describing the clustering of collapsed objects across distances that approach the Hubble distance, and the propagation of light from very distant and early objects over a time scale comparable to the Hubble time. What is more, even a small object could have formed out of primordial density perturbation across a super-Hubble patch at early times. Therefore, a general relativistic understanding of nonlinear large-scale structure formation is desirable.

It is noticable that in many contexts the problem boils down to how short-scale physics, gravitational or baryonic, may be systematically affected by a background perturbation spanning much larger scales. Such a scale hierarchy exists, for instance, in the study of how the collapse of dark matter halos or galaxies out of small lumps of primordial overdensity may be modulated by their surrounding mean environment, which varies very slowly across large distances. This results in the phenomenon of biased clustering of clusters, galaxies, and so on, with respect to the dark matter distribution on large scales.[198,199] In a similar way, large-scale gravitationial tidal field can have a systematic effect on the dark matter halo spin or the galaxy orientation, giving rise to the phenomenon of intrinsic shape alignment. In many cases, such systematic influence from large-scale perturbations lead to quantitative predictions with clean physical interpretation.

An intuitive way to analyze how long-wavelength perturbations can affect short-scale physics is to conjecture that the short-scale physics is *not* sensitive to the spatial variation of the long-wavelength perturbations, so that the effect is equivalent to a local modification of cosmology. This means that the average matter density Ω_m (and the spatial geometry), the expansion rate H, and so on are locally modified. This is the perspective of *separate*

universes pioneered by previous analytical and computational studies.[198, 200–204]

A simple, vivid example for the separate universes approach is the simple case of spherical top-hat collapse. Consider a homogeneous flat FLRW universe, with an average matter density $\bar{\rho}(t)$ satisfying the Friedmann equation. Now imagine it is perturbed by a top-hat overdensity $\rho(t) > \bar{\rho}(t)$, localized within some finite radius but is perfectly spherical symmetric and uniform. Due to *Birkhoff's theorem*,[205] the dynamics in the interior of the spherical top-hat disregards the matter distribution in the exterior, i.e. the global homogeneous FLRW universe. Then it is as if the top-hat is a different homogeneous universe, with a uniform matter density $\rho(t)$ higher than the critical value $\bar{\rho}(t)$, and therefore it must obey a different Friedmann equation for curved, over-critical FLRW cosmology. The same thing will also be true for an underdense region, and irrespective of the size of the density contrast with respect to the global $\bar{\rho}(t)$. Furthermore, due to exact spherical symmetry, the picture is precise in full general relativity, regardless of the radius of the top-hat. It is even possible to have a number of top-hat over-/under-densities of different sizes at different locations, each of which would evolve as a separate FLRW universe, making it a "Swiss-Cheese" universe.[206–208] Besides, the conclusion can be generalized to include a finite cosmological constant.

With a stochastical initial perturbation upon a homogeneous FLRW universe, the surrounding environment for a local structure formation site will not be perfectly spherical symmetric, and therefore the aforementioned spherical top-hat picture do not exactly apply. Still, the departure can be locally captured by a tidal shear, which characterizes the leading anisotropic effect from long-wavelength perturbations on short-scale physics. Addi-

tional corrections are anticipated to scale with the length ratio between the short-wavelength physics and long-wavelength physics, and are practically negligible in many cases.

In this Chapter, we aim at developing a rigorous theory of the separate universe approach within the framework of general relativity. Lagrangian observers who are freely falling in a smoothly perturbed Universe (up to small-scale perturbations and peculiar velocity) define a natural frame to interpret the evolution and geometry of the Universe in their vicinity. If such an observer is incapable of distinguishing between the local effect of a long-wavelength perturbation and a modified cosmology by conducting any local measurement, it strongly suggests that a local coordinate system equivalent to the modified FLRW can be constructed. This leads to the concept of *Conformal Fermi Coordinates* (CFC hereafter),[174] a cosmological generalization of the famous *Fermi Normal Coordinates* (FNC hereafter).[209] The CFC has the advantage that gauge artifacts encountered in conventional cosmological perturbation theory are unambigiously fixed by reference to proper time and proper length, and it allows for a systematic derivation of the local gravitational effect from a general long-wavelength perturbation. Using the CFC formalism we shall draw important conclusion on the phenomenology of large-scale structure, including the squeezed galaxy bispectrum and scale-dependent galaxy bias.

This Chapter will be heavily based on Ref.[210] and Ref.[162] We will begin with a presentation on the construction of the CFC in Section 6.1. We then discuss in Section 6.2 the general form of Einstein equations and fluid equations in the CFC frame that correctly capture the physical nonlinear coupling between long-wavelength modes and short-wavelength modes. After that, we apply the CFC formalism to the case of long-wavelength

scalar perturbation in Section 6.3, and then present a computation of the squeezed-limit matter bispectrum in Section 6.4. In the end, concluding remarks will be given in Section 6.5.

A note on our notations: the wave vector of a long-wavelength perturbation will be denoted by $\mathbf{k}_L$, and that of a short-wavelength perturbation will be denoted by $\mathbf{k}_S$.

6.1 Conformal Fermi Coordinates

We have argued from physical intuition that the local effects of a long-wavelength perturbation, to leading order, mimick that of a homogeneous, but modified FLRW cosmology. Therefore, there must exist a natural coordinate system for a free-falling local observer, whose metric has approximately the FLRW form. This is called the Conformal Fermi Coordinates (CFC), a direct generalization of the Fermi Normal Coordinates (FNC).[209]

We would like to clarify that the local CFC is constructed for a smoothed spacetime with only long-wavelength perturbations. This can be achieved by first filtering out the ubiquitous fluctuations on small scales in the Universe. Through this procedure of coarse graining, we are able to construct a slowly-varying background spacetime, in which we then study the dynamics of short-scale physics.

We now describe the construction of the CFC. The subscript F will be reserved for CFC hereafter. The CFC are constructed in the vicinity of a *central geodesic* G, usually taken to be the trajectory of a free-falling observer. In CFC it is required that the lowest

order CFC metric has an FLRW form,

$$g^F_{\mu\nu}(x^\mu_F) = a^2_F(\tau_F)\left[-\eta_{\mu\nu} + h^F_{\mu\nu}(\tau_F, x^i_F)\right], \qquad h^F_{\mu\nu} = \mathcal{O}[(x^i_F)^2]. \tag{6.1}$$

Let x^i_F be the spatial displacement from the CFC spatial origin. Correction $h^F_{\mu\nu}$ to the FLRW form is required to start at quadratic order in x^i_F, which can be called the tidal metric correction. The CFC time τ_F should be some suitable conformal time rather than the observer's proper time. Besides, a suitable local scale factor $a_F(\tau_F)$ should be defined in a physical way. These considerations motivate our construction of the CFC as presented below.

6.1.1 Construction of CFC

Figure 6.1 shows the geometrical relation between the CFC and the global coordinates. Throughout this paper, "global coordinates" refers to some standard coordinate system valid in a region surrounding the central geodesic G, to be distinguished from the locally constructed CFC. This includes any standard gauge to parametrize metric and matter perturbations.

First, we choose a set of orthonormal tetrad vectors: the CFC temporal direction $(e^\mu_0)_P$ is along the tangent direction of G, and the space-like $(e^\mu_i)_P$, $i = 1, 2, 3$ are three spatial directions perpendicular to the temporal direction. The observer's geodesic G defines the spatial origin $x^i_F = 0$ in the CFC and is parametrrized by the proper time t_F.

For a suitably chosen local expansion $a_F(t_F)$ along G, we may define a "conformal

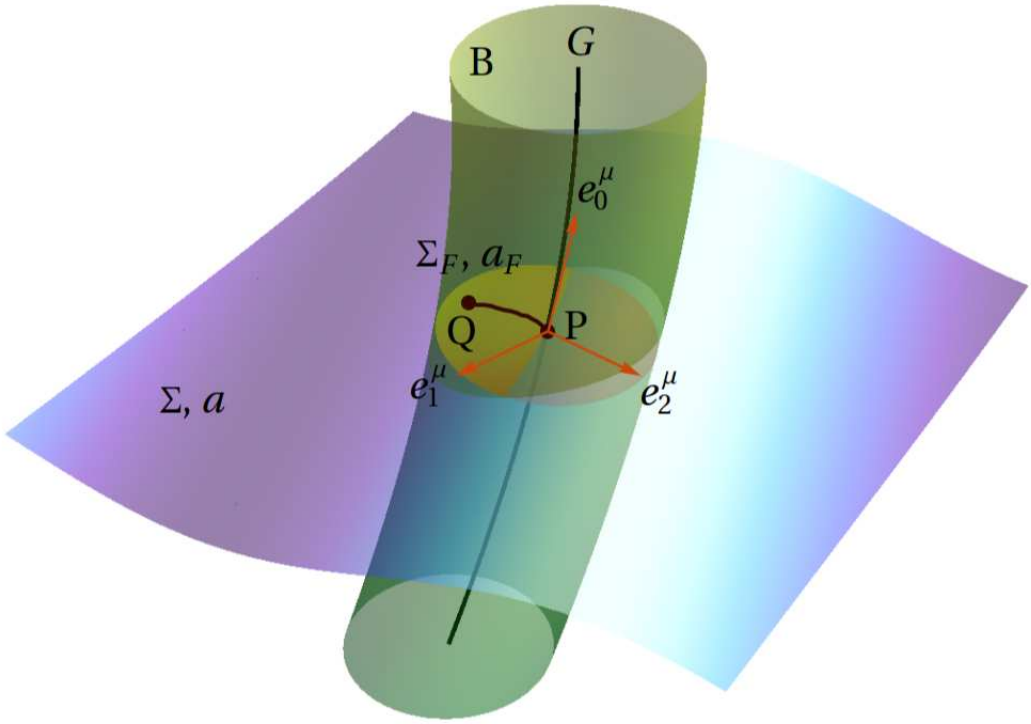


Figure 6.1: Construction of the CFC from Ref.[210] At point P, the observer's geodesic G intersects a spatial hypersurface Σ, having constant conformal time τ and scale factor $a(\tau)$ in some global coordinate system. The spatial hyper-surface Σ_F, having constant CFC conformal time τ_F and CFC scale factor $a_F(\tau_F)$, also intersects G at P, but does not coincide with Σ in general. Any other point Q on Σ_F may be connected to P by a conformal geodesic, and is parameterized by the same τ_F but nonzero x_F^i. The CFC coordinates are valid within a tubelike region bounded by hyper-surface B surrounding G.

proper time" τ_F through

$$d\tau_F = a_F^{-1}(P(t_F))dt_F\,, \tag{6.2}$$

where $P(t_F)$ is the point along the central geodesic at proper time t_F. This can be integrated to yield a unique relation $\tau_F(t_F)$, up to an integration constant which can be absorbed into a redifinition of a_F. We then choose τ_F as the CFC time coordinate.

We then define the slice of constant τ_F to be orthogonal to G where both intersect

with each other. In a homogeneous FLRW universe, an arbitrary point Q on a certain slice of constant time can be connected to the spatial origin of that slice with "straight lines". However, those are not true geodesics of the expanding spacetime. Instead, they would be true geodesics if the scale factor $a^2(\tau)$ is removed from the FLRW metric, i.e. the flat metric $\eta_{\mu\nu} = a^{-2} g_{\mu\nu}$. This provides a hint that surfaces of constant-τ_F in CFC can be defined to be spanned by space-like *conformal geodesics*, namely geodesics with respect to the conformal metric i.e. .

$$\tilde{g}_{\mu\nu}(x) \equiv a_F^{-2}(x) g_{\mu\nu}(x)\,. \tag{6.3}$$

It is crucial to note that for a perturbed FLRW metric $g_{\mu\nu} = a^2(\eta_{\mu\nu} + h_{\mu\nu})$, $\tilde{g}_{\mu\nu}$ in general differs from $\eta_{\mu\nu} + h_{\mu\nu}$ because the local scale factor a_F is not necessarily identical to the global one a. Here and throughout, a tilde will be used to denote quantities defined with respect to this conformal metric.

The conformal geodesic connecting Q and P satisfies the conformal geodesic equation

$$\frac{d^2 x^\mu}{d\lambda^2} + \tilde{\Gamma}^\mu_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0\,, \tag{6.4}$$

Here the Christoffel symbols for the metric $g_{\mu\nu}$ are replaced by the conformally transformed ones

$$\tilde{\Gamma}^\mu_{\alpha\beta} = \Gamma^\mu_{\alpha\beta} - C^\mu{}_{\alpha\beta}\,. \tag{6.5}$$

Here we define the tensor[211]

$$C^{\mu}{}_{\alpha\beta} = \delta^{\mu}_{\alpha}\,\nabla_{\beta}\ln a_F + \delta^{\mu}_{\beta}\,\nabla_{\alpha}\ln a_F - g_{\alpha\beta}\,g^{\mu\lambda}\nabla_{\lambda}\ln a_F\,. \tag{6.6}$$

In order to obtain the leading correction to the CFC metric, it is sufficient to specify the first and the second derivatives of a_F. The first derivative of a_F is fixed by Eq. (6.1). In order to have the CFC form, the gradient of a_F has to be along the tangent direction of the central geodesic G, i.e.

$$\nabla_{\mu}\ln a_F\Big|_{x^i_F=0} = (\ln a_F)'\,a_F\,(e_0)_{\mu}\,, \tag{6.7}$$

where a prime $'$ denotes a derivative with respect to τ_F. The second derivative $\nabla_{\mu}\nabla_{\nu}\ln a_F|_{x^i_F=0}$ is chosen to precisely match the relation in an unperturbed FLRW universe

$$\nabla_{\alpha}\nabla_{\beta}\ln a_F\Big|_{x^i_F=0} = \left[\frac{1}{a_F^2}\frac{d\mathcal{H}_F}{d\tau_F} - 2\left(\frac{\mathcal{H}_F}{a_F}\right)^2\right](e_0)_{\alpha}(e_0)_{\beta} - \left(\frac{\mathcal{H}_F}{a_F}\right)^2 g_{\alpha\beta}. \tag{6.8}$$

To define the spatial CFC coordinate x^i_F for Q, let us consider the Taylor expansion of Eq. (6.4) with respect to the affine parameter

$$x^{\mu}(\lambda) = \sum_{n=0}^{\infty}\alpha^{\mu}_n\lambda^n\,. \tag{6.9}$$

This curve connects point P ($\lambda = 0$) at CFC location $x^{\mu}_F(P) = \{\tau_F, \mathbf{0}\}$ with point Q ($\lambda = 1$) at CFC location $x^{\mu}_F(Q) = \{\tau_F, x^i_F\}$. Since P is chosen to be the spatial origin, $\alpha^{\mu}_0 = x^{\mu}(P)$. As the next coefficient in the expansion, the tangent vector of the conformal geodesic at

$\lambda = 0$ gives the definition of x^i_F through

$$\alpha^\mu_1 = \left.\frac{dx^\mu}{d\lambda}\right|_{\lambda=0} = a_F(P)(e_i)^\mu_P\, x^i_F\,, \tag{6.10}$$

It is defined such that to leading order $a_F(\tau_F)x^i_F$ gives the proper distance from Q to P. In principle, higher-order coefficients α^μ_n can be then recursively obtained using Eq. (6.4). For instance, the second-order coefficient is given by

$$\alpha^\mu_2 = \frac{1}{2}\left.\frac{d^2x^\mu}{d\lambda^2}\right|_P = -\frac{1}{2}\,\tilde{\Gamma}^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\bigg|_P = -\frac{1}{2}a^2_F(P)\left(\tilde{\Gamma}^\mu_{\alpha\beta}\right)_P (e_i)^\alpha_P (e_j)^\beta_P\, x^i_F x^j_F, \tag{6.11}$$

and the third-order coefficient is

$$\begin{aligned}\alpha^\mu_3 &= \frac{1}{6}\left.\frac{d^3x^\mu}{d\lambda^3}\right|_P = -\frac{1}{6}\frac{d}{d\lambda}\left(\tilde{\Gamma}^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\right)_P \\ &= -\frac{1}{6}a^3_F(P)\left(\partial_\gamma\tilde{\Gamma}^\mu_{\alpha\beta} - 2\tilde{\Gamma}^\mu_{\sigma\alpha}\tilde{\Gamma}^\sigma_{\beta\gamma}\right)_P (e_i)^\alpha_P\,(e_j)^\beta_P\,(e_k)^\gamma_P\; x^i_F x^j_F x^k_F.\end{aligned} \tag{6.12}$$

These expressions are useful to derive the coordinate transformation from any global coordinate system to the CFC and how physical quantities transform accordingly.

6.1.2 CFC metric

We now derive the CFC metric Eq. (6.1). We in particular determine the quadratic terms, which represent the leading departure from a homogeneous FLRW spacetime. Instead of referring to the original global coordinate system, this task can be most easily and intuitively done by thinking in terms of the CFC coordinates directly, and projecting

various geometric quantities into the CFC frame.

To begin with, the CFC temporal coordinate vector is $(\tilde{e}_0) \equiv \partial/\partial x^0_F = a_F(P)(e_0)_P$, and the spatial coordinate vectors are $(\tilde{e}_i) \equiv \partial/\partial x^i_F = a_F(P)(e_i)_P$. Because those are orthonormal, the CFC metric evaluated right on the central geodesic must be conformally flat,

$$g^F_{\mu\nu}\big|_P = \Big[(\tilde{e}_\mu)^\alpha\,(\tilde{e}_\nu)^\beta\,g_{\alpha\beta}\Big]_P = a^2_F(P)\,\Big[(e_\mu)^\alpha\,(e_\nu)^\beta\,g_{\alpha\beta}\Big]_P = a^2_F(P)\,\eta_{\mu\nu}, \tag{6.13}$$

According to our construction, the spatial CFC coordinate lines, being geodesics with respect to the conformal metric, are simply parametrized in the CFC through

$$x^0_F = \tau_F = \text{const.}, \qquad x^i_F = \beta^i\lambda\,, \tag{6.14}$$

where β^i is a constant vector and λ is some affine parameter. Since the tangent vector is given by $\{0, \beta^i)\}$, in the CFC frame the conformal geodesic equation Eq. (6.4) simplifies to

$$(\tilde{\Gamma}^F)^\mu_{ij}\Big|_P\,\beta^i\beta^j = 0, \tag{6.15}$$

which implies $(\tilde{\Gamma}^F)^\mu_{ij}|_P = 0$ because β^i is arbitrary. One can further find that following Eq. (6.14), all Christoffel symbols $(\tilde{\Gamma}^F)^\mu_{\alpha\beta}$, computed with respect to $\tilde{g}_{\mu\nu}$ [Eq. (6.3)] and projected into the CFC frame, vanish on the central geodesic, i.e. $(\tilde{\Gamma}^F)^\mu_{\alpha\beta}|_P = 0$. It further follows that all first derivatives of the conformal metric $\tilde{g}_{\mu\nu}$ in CFC vanish along the central

geodesic,

$$(\partial_\alpha \tilde{g}_{\mu\nu})^F\Big|_P = 0\,. \tag{6.16}$$

This gaurantees the absence of term of $\mathcal{O}[x^i_F]$ in the CFC metric Eq. (6.1).

Furthermore, the $\mathcal{O}[(x^i_F)^2]$ terms in the CFC metric are found to be related to the *conformal* Riemann curvature tensor $\tilde{R}$, constructed for the conformal metric, and evaluated on G. Paralleling the analysis presented in Ref.,[209] the quadratic corrections can be derived from the equation of geodesic deviation. In the end we obtain

$$g^F_{00}(x_F) = a^2_F(\tau_F)\left[-1 - \tilde{R}^F_{0k0l}\Big|_P x^k_F x^l_F\right], \tag{6.17}$$

$$g^F_{0i}(x_F) = a^2_F(\tau_F)\left[-\frac{2}{3}\,\tilde{R}^F_{0kil}\Big|_P x^k_F x^l_F\right], \tag{6.18}$$

$$g^F_{ij}(x_F) = a^2_F(\tau_F)\left[\delta_{ij} - \frac{1}{3}\,\tilde{R}^F_{ikjl}\Big|_P x^k_F x^l_F\right]. \tag{6.19}$$

Here $\tilde{R}^F$ is the Riemann curvature tensor constructed with respect to $\tilde{g}_{\mu\nu}$ and has its components projected in the CFC frame. In terms of some global coordinates,

$$\tilde{R}^F_{\alpha\beta\gamma\delta} = (\tilde{e}_\alpha)^\mu(\tilde{e}_\beta)^\nu(\tilde{e}_\gamma)^\rho(\tilde{e}_\delta)^\sigma \tilde{R}_{\mu\nu\rho\sigma}, \tag{6.20}$$

where the covariant Riemann curvature tensor $\tilde{R}_{\mu\nu\rho\sigma}$ has its components computed in any global coordinates.

The extension of this analysis to higher-order metric correction terms $\mathcal{O}[(x^i_F)^n]$ in CFC can become increasingly sophisticated. The bottom line, however, is that those terms are

order-by-order suppressed by the spatial derivatives of the long-wavelength perturbation. If we focus on the vicinity of the CFC origin $x^i_F = 0$, those higher-order corrections are parametrically suppressed.

6.1.3 CFC scale factor

We have introduced the concept of a local scale factor $a_F(\tau_F)$, which describes how fast the local space around the CFC observer expands as a function of the local time. However, we have not specified how the local scale factor can be determined. In fact, following the aforementioned construction of CFC frame, the form Eq. (6.1) is preserved under a general reparametrization of the CFC conformal time

$$\tau_F = \tau_F(\tau'_F), \qquad a_F(\tau_F) = a'_F(\tau'_F)\,\frac{d\tau'_F}{d\tau_F}. \tag{6.21}$$

Note that the observer's proper time $a_F(\tau_F)d\tau_F = dt_F$ is also invariant under the reparametrization. Of course, the new $x^{i\prime}_F$'s have to be re-defined accordingly, following our construction with the conformal geodesics. In general, this would lead to a different choice for constant-time slicing, and result in different CFC metric perturbations Eqs. (6.17)–(6.19). This suggests that without a given choice of $a_F(\tau_F)$, the aforementioned construction of CFC is not unique. As an extreme but simple case, $a_F = 1$ would correspond to the locally flat FNC.

It is physically favorable if a_F is chosen to be a locally observable quantity with physical significance. In a homogeneous FLRW universe, "expansion of space" can be operationally

measured from the systematic Hubble flow of nearby astronomical objects. Similarly, in the "local universe", we may measure a local Hubble flow. Imagine that many free-falling observers move with 4-velocity U^μ in the neighborhood of the CFC observer. Then the expansion of space can be quantified by the *divergence of geodesic congruence*

$$H_F = \frac{d\ln a_F}{dt_F} = \nabla_\mu U^\mu = \frac{1}{3}\vartheta \tag{6.22}$$

Indeed, this choice reduces to the global expansion rate when metric perturbation to an FLRW universe is set to zero. Eq. (6.22) in fact determines a_F up to a multiplicative constant, which fixes the normalization of comoving scales.

We pass by noting that in Appendix E there are still residual gauge freedom in the spatial CFC metric h^F_{ij} at $\mathcal{O}[(x^i_F)^2]$, even when a prescription has been specified to fix the local scale factor $a_F(\tau_F)$.[174]

In some global coordinate system, the coarse-grained universe can be described by an FLRW spacetime perturbed by metric perturbation

$$ds^2 = a^2(\tau)\left[\eta_{\mu\nu} + h_{\mu\nu}\right] dx^\mu dx^\nu \,. \tag{6.23}$$

We stress again that in Eq. (6.23) small-scale fluctuations have been filtered out. To stay fully general, we do not make a specific gauge choice for $h_{\mu\nu}$. For practical application, in the following we will assume that $h_{\mu\nu}$ is numerically small, so that it suffices to stay accurate at linear order in h. It is worth to note, however, that the perturbative expansion

in $h_{\mu\nu}$ should be distinguished from the power expansion in x^i_F. The former expansion is applicable as long as $|h_{\mu\nu}| \ll 1$, and is adopted here for the sake of calculational simplicity. On the other hand, the latter expansion is essentially an expansion in the spatial derivatives of the metric perturbation, and it holds if $|x^i_F|$ is smaller than the typical scale of variation of $h_{\mu\nu}$. This defines the regime of validity for the CFC frame.

Up to linear order in metric perturbations, the observer's 4-velocity can be parameterised as $U^\mu = a^{-1}\left(1 + h_{00}/2, V^i_o\right)$, where the 3-velocity V^i_o is to be treated as first-order perturbation. The tetrad vectors have explicit forms

$$(e_0)^\mu = a^{-1}\left(1 + \frac{1}{2}h_{00}, V^i_o\right), \tag{6.24}$$

$$(e_i)^\mu = a^{-1}\left(V_{o,j} + h_{0j}, \delta^i{}_j - \frac{1}{2}h^i{}_j + \frac{1}{2}\varepsilon_j{}^{ik}\omega_k\right), \tag{6.25}$$

In the presence of vector metric perturbations, it is consistent to introduce the rotation vector ω^i. Since the tetrad vectors are parallel transported along the central geodesic, V^i_o and ω^i satisfy the following equations of motion, respectively,

$$V^{i\prime}_o + \mathcal{H}V^i_o = \frac{1}{2}\partial^i h_{00} - h^i{}_0{}' - \mathcal{H}h^i{}_0, \tag{6.26}$$

$$\omega^{k\prime} = -\frac{1}{2}\varepsilon^{kij}\left(\partial_i h_{0j} - \partial_j h_{0i}\right). \tag{6.27}$$

Here a prime denotes the derivative with respect to the conformal time.

We find the divergence of geodesic congruence ϑ to be given by

$$\vartheta = \frac{1}{a}\left[3\mathcal{H} + \frac{3}{2}\mathcal{H}h_{00} + \frac{1}{2}h' + \partial\cdot V_o\right], \tag{6.28}$$

where $h = \delta^{ij}h_{ij}$. On the right hand side, it is to remember that all quantities are to be computed in global coordinates but along the CFC central geodesic. The local expansion rate is then given by the local Hubble parameter

$$\frac{\mathcal{H}_F(\tau_F)}{a_F(\tau_F)} = \frac{1}{a(\tau)}\left[\mathcal{H}(\tau) + \frac{1}{2}\mathcal{H}(\tau)h_{00} + \frac{1}{6}h' + \frac{1}{3}\partial\cdot V_o\right]. \tag{6.29}$$

The normalization of the CFC scale factor can be fixed by the requirement that a_F coincides with a when evaluated at some early, but *fixed* CFC proper time t_F,

$$\frac{a_F(\tau_F(t_F))}{a(\tau(t=t_F))} \to 1, \qquad \text{as } t_F \to 0\,. \tag{6.30}$$

In a matter-dominated universe, this is true if we choose

$$\frac{a_F(\tau_F)}{a(\tau)} \to 1 - \frac{1}{3}\,h_{00,\mathrm{ini}}, \qquad \text{as } \tau_F \to 0\,. \tag{6.31}$$

where $h_{00,\mathrm{ini}} = h_{00}(\tau_{\mathrm{ini}})$ is the metric perturbation evaluated on the geodesic when perturbation modes are on superhorizon scales. Furthermore, Eq. (6.29) yields

$$\frac{d}{d\tau}\frac{a_F(\tau_F)}{a(\tau)} = -\frac{a_F\mathcal{H}}{a} + \frac{a_F\mathcal{H}_F}{a}\frac{d\tau_F}{d\tau} = \frac{a_F}{a}\left[\frac{1}{6}h' + \frac{1}{3}\partial\cdot V_o\right], \tag{6.32}$$

which can be integrated to give

$$\frac{a_F(\tau_F)}{a(\tau)} = 1 - \frac{1}{3}H\, h_{00,\mathrm{ini}} + \int_{\tau_{\mathrm{ini}}}^{\tau} d\tau \left(\frac{1}{6}h' + \frac{1}{3}\partial\cdot V_o\right) . \tag{6.33}$$

This defines the physical, locally measurable CFC scale factor for the most general FLRW metric with linear long-wavelength perturbations.

6.2 Einstein Equations and Fluid Equations

After constructing the CFC for the coarse-grained universe, we need to put small-scale perturbations back into Eq. (6.1) in order to study their evolution in this local frame. In particular, the small-scale perturbations obey a set of Einstein equations (the fundamental equation of gravitational dynamics) and fluid equations (the consequence of energy-stress conservation). The appropriate form of those equations for short-scale perturbations are non-trivial in CFC, as they must correctly represent the local, gauge invariant influence of long-wavelength perturbations on short-scale physics. To account for the nonlinear couplings between long-wavelength perturbations and short-wavelength perturbations, we must at least work to second order in perturbation, which is often a formidable task. As we will see, special properties of the CFC frame can result in substantial simplification in those equations. We begin with the discussion on Einstein equations.

6.2.1 Coarse Graining

In reality, spacetime metric is inhomogeneous down to very small scales. For the CFC metric to be valid over a finite region, we introduce a coarse-graining of the metric on some comoving spatial scale $L = \Lambda^{-1}$ to allow for the CFC metric to be valid across a region of sufficient size surrounding the central geodesic. The CFC metric is then constructed with respect to the coarse-grained metric perturbation h^{Λ} with wavelengths in Fourier space $k^{-1} \gtrsim \Lambda^{-1}$. In this section, we describe the coarse-graining and derive the structure of Einstein equations in the CFC frame. Throughout we will keep up with a perturbative calculation at linear order in the coarse-grained metric perturbation, but in principle allow the small-scale perturbations to be nonlinear.

In some global coordinate system the Einstein equations read

$$G_{\mu\nu}[g] = 8\pi G\, T_{\mu\nu}\,. \tag{6.34}$$

Subtracting the homogeneous background solution, we obtain the equation for the perturbation part $h_{\mu\nu}$:

$$G^{a(\tau)}_{\mu\nu}[h] = 8\pi G\, T_{\mu\nu}\,, \tag{6.35}$$

where $G^{a(\tau)}[h]$ is the perturbation to the Einstein tensor, which depends on $a(\tau)$, and $T_{\mu\nu}$ only contains the perturbation to the stress energy tensor. Since we no longer need the homogeneous background stress energy tensor, we have kept the same symbol $T_{\mu\nu}$ for simplicity.

We can apply the coarse-graining procedure to the stress-energy tensor and to the metric

$$h_{\mu\nu} = h^{\Lambda}_{\mu\nu} + h^{s}_{\mu\nu}; \quad T_{\mu\nu} = T^{\Lambda}_{\mu\nu} + T^{s}_{\mu\nu}\,, \tag{6.36}$$

where the "s" components are defined as the difference between the precise quantity and its coarse-grained counterpart, representing short-scale fluctuations. We may further split the Einstein tensor into linear and nonlinear pieces in the metric perturbation, $G^{a(\tau)}[h] = G^{(\mathrm{lin})}[h] + G^{(\mathrm{nl})}[h]$, where the linear piece $G^{(\mathrm{lin})}$ apparently commutes with coarse graining.

We can then coarse-grain Eq. (6.35) to find

$$G^{(\mathrm{lin})}_{\mu\nu}[h^{\Lambda}] + G^{(\mathrm{nl})\Lambda}_{\mu\nu}[h] = 8\pi G\, T^{\Lambda}_{\mu\nu}\,, \tag{6.37}$$

which is equivalent to an implicit equation for h^{Λ},

$$G^{(\mathrm{lin})}_{\mu\nu}[h^{\Lambda}] + G^{(\mathrm{nl})}_{\mu\nu}[h^{\Lambda}] = 8\pi G\, T^{\Lambda}_{\mu\nu} + \left\{ G^{(\mathrm{nl})}_{\mu\nu}[h^{\Lambda}] - G^{(\mathrm{nl})\Lambda}_{\mu\nu}[h] \right\}\,. \tag{6.38}$$

The effective source in curly brackets on the right hand side accounts for the backreaction of small scales perturbations to the dynamics on large scales. In reality, backreaction is a small effect, which we will ignore in the following.

Subtracting Eq. (6.38) from Eq. (6.35), and dropping the additional source, we derive an equation for h^{s}:

$$G^{(\mathrm{lin})}_{\mu\nu}[h^{s}] + G^{(\mathrm{nl})}_{\mu\nu}[h] - G^{(\mathrm{nl})}_{\mu\nu}[h^{\Lambda}] = 8\pi G\, T^{s}_{\mu\nu}\,, \tag{6.39}$$

To quantify how the long mode perturbatively affects the small-scale fluctuation h^s, we power expand in a series in h^Λ:

$$h^s = h^{s0} + h^{s\Lambda} + h^{s\Lambda^2} + \cdots , \tag{6.40}$$

where $h^{s\Lambda^n}$ scales as $(h^\Lambda)^n$. For example, h^{s0} would be the small-scale perturbation in the absence of the long-wavelength modes $h^\Lambda \to 0$, and $h^{s\Lambda}$ captures the linear response to the long modes. Not going to nonlinear orders in h^Λ, just for simplicity, we will truncate the series at $h^{s\Lambda}$ in the following. Note that T^s is defined through Eq. (6.36). Similar expansion can be applied to the energy-stress tensor $T^\mu{}_\nu$ At zeroth order in the long mode, the Einstein equation

$$G^{(\rm lin)}_{\mu\nu}[h^{s0}] + G^{(\rm nl)}_{\mu\nu}[h^{s0}] = 8\pi G\, T^{s0}_{\mu\nu} , \tag{6.41}$$

just reduces to the linear perturbation theory for small-scale fluctuations. At linear order in h^Λ we obtain the equation for $h^{s\Lambda}$:

$$G^{(\rm lin)}_{\mu\nu}[h^{s\Lambda}] + G^{(\rm nl)\prime}_{\mu\nu}[h^{s0}, h^\Lambda + h^{s\Lambda}] = 8\pi G\, T^{s\Lambda}_{\mu\nu} , \tag{6.42}$$

with the "response" operator

$$G^{(\rm nl)\prime}[h, \tilde{h}] = \frac{\partial G^{(\rm nl)}[h]}{\partial h_{\mu\nu}}\tilde{h}_{\mu\nu} + \frac{\partial G^{(\rm nl)}[h]}{\partial(\partial_\alpha h_{\mu\nu})}\partial_\alpha\tilde{h}_{\mu\nu} + \frac{\partial G^{(\rm nl)}[h]}{\partial(\partial_\alpha\partial_\beta h_{\mu\nu})}\partial_\alpha\partial_\beta\tilde{h}_{\mu\nu} . \tag{6.43}$$

It is desirable that Eq. (6.42) is linear in the long-wavelength perturbation, but in general fully nonlinear in h^{s0}. This equation gives the leading impact of long-wavelength perturba-

tions on the dynamics of short-scale physics.

Our ultimate goal is to derive the same equation in the CFC frame. In this case, however, long-wavelength metric perturbations enter only through a modification of the expansion rate, or through the tidal metric corrections $h^{\Lambda,F}_{\mu\nu}$ at $\mathcal{O}[(x^i_F)^2]$. We have the counterpart of Eq. (6.43) in the CFC

$$G^{(\mathrm{lin})\mathrm{F}}[h^{s\Lambda}_F] + G^{(\mathrm{nl})\mathrm{F}\prime}[h^{s0}, h^{s\Lambda}_F] + G^{(\mathrm{nl})\mathrm{F}\prime}[h^{s0}, h^{\Lambda}_F] = 8\pi G\, T^{s\Lambda,F}\,, \tag{6.44}$$

where the tensor indices have been suppressed just for simplicity. Note that the tensors $G^{(\mathrm{lin})\mathrm{F}}_{\mu\nu}$ and $G^{(\mathrm{nl})\mathrm{F}\prime}_{\mu\nu}$ are computed with respect to the CFC metric. In particular, they involve the local scale factor a_F rather than the background one.

6.2.2 Einstein Equations

In the following, we shall omit the "$_F$" label everywhere, as we will only be dealing with the CFC metric, except in h^{Λ}_F. The equations simplify if we aim to solve the Einstein equation *on the central geodesic* $\mathbf{x}_F = 0$. As we will see later, this turns out to be sufficient because we are free to construct the CFC around any timelike geodesic.

Keeping up to second order in perturbation, we drop the second term on the left hand side of Eq. (6.44) as it is of higher order. For the third term on the left hand side, we need at least two spatial derivatives acting on h^{Λ}_F to obtain a nonzero contribution at $x^i = 0$. Since the nonlinear Eintein tensor contains at most two spacetime derivatives acting on metric perturbation, h^{s0} must enter without any spatial derivative. Therefore, we can write

Eq. (6.44) as

$$\text{CFC, 2nd order:} \quad G^{(\text{lin})}_{\mu\nu}[h^{s\Lambda}] + \mathcal{O}(h^{s0}\partial_i\partial_j h^{\Lambda}_F)_{\mu\nu} = 8\pi G T^{s\Lambda}_{\mu\nu} \,. \tag{6.45}$$

We thus only have to work out a limited number of terms.

On the other hand, in the subhorizon limit for the small-scale perturbations h^s, the leading terms have double spatial derivatives acting on the metric perturbation. On the other hand, we allow the long-wavelength perturbation h^Λ to be arbitrarily long. In the subhorizon limit, Eq. (6.44) reduces to

$$\text{CFC, subhorizon:} \quad G^{(\text{lin,SH})}_{\mu\nu}[h^{s\Lambda}] = 8\pi G T^{s\Lambda}_{\mu\nu} \,, \tag{6.46}$$

where $G^{(\text{lin,SH})}_{\mu\nu}$ is the linearized Einstein tensor around the *local* FLRW background $a_F(\tau_F)$. We therefore conclude that *the long-wavelength modes do not appear on the left hand side of the CFC frame Einstein equations in the subhorizon limit*, which would not be the case in global coordinates, where terms of the form $h^\Lambda \partial\partial h^{s0}$ are in general present.

On the right hand side of the Einstein equations, the metric enters the energy-stress tensor with no spatial derivatives, so that h^Λ does not appear at all when evaluated at $x^i = 0$.

6.2.3 Fluid Equations

To close the system of equations, we need to supply with the fluid equations, which follow from energy-stress conservation

$$\nabla_\mu T^{\mu\nu} = 0\,, \tag{6.47}$$

In the same manner as described above, we can derive the fluid equation that describes the evolution of the long-short mode coupling contribution $T^{s\Lambda}$. Naively, h_F^Λ enters this equation with only one spatial derivative, leading to a vanishing contribution on the central geodesic in CFC. However, spatial derivatives of the velocity $\partial_i v^j \sim \partial_i T^j{}_0$ appear in the fluid equations. Thus, in order to close the hierarchy, we need to take one further spatial derivative. Since the velocity itself depends on ∂h^Λ, the equation for $\partial_i v^i$ involves two spatial derivatives of the metric. In this way, h_F^Λ enters in the small-scale dynamics in CFC through the fluid equations. Another way to say this is that we need to know the velocity at order x_F away from the central geodesic in order to derive the evolution of the density on the central geodesic.

6.3 Long-wavelength scalar perturbations and CFC

In this Section, we study long-wavelength scalar metric perturbations using the CFC formalism we have developed. Those are the gravitational potential generated by the density perturbation in the early Universe. Scalar perturbations exist in the initial condition and

span an almost scale-invariant spectrum, so that there are both modes having very long wavelengths and modes having short wavelengths, and the evolution of the latter is coupled to the former because the process of gravitational clustering is nonlinear. In this section, we study this coupling in the CFC frame.

We begin with the following coarse-grained metric with long-wavelength modes

$$ds^2 = a^2(\tau)\left[-(1+2\Phi)d\tau^2 + (1-2\Psi)\delta_{ij}dx^i dx^j\right] . \tag{6.48}$$

This is in the conformal Newtonian (cN) gauge, and the perturbation variables are the two gravitational potentials Φ and Ψ. Note that the cN gauge is merely a convenient choice; other gauges may be used to define the same CFC, since the construction of Section 6.1.1 is coordinate independent. In the following, we shall assume that Φ and Ψ are small and only compute to linear order.

A free-falling observer in the coarse-grained universe has her 4-velocity given by

$$U^\mu = (e_0)^\mu = a^{-1}\left(1-\Phi, V^i\right) , \tag{6.49}$$

where the 3-velocity V^i is considered as first-order perturbation.[1] V^i obeys the equation for geodesic motion

$$V^{i\prime} + \mathcal{H}V^i \;\; = \;\; -\partial^i\Phi\,. \tag{6.50}$$

[1]Throughout, uppercase V^i is reserved for large-scale peculiar velocity. Later we will introduce the small-scale peculiar velocity denoted by lowercase v^i. Similar notations apply to other perturbation variables such as the density field.

Eq. (6.49) gives the observer's tetrad vector in the temporal direction. The spatial tetrad vectors orthogonal to it may be chosen as

$$(e_j)^\mu \;=\; a^{-1}\left(V_j, \delta^i{}_j[1+\Psi]\right)\,, \tag{6.51}$$

Computing the divergence of geodesic congruence as in Eq. (6.29), the local expansion rate is given by

$$H_F = H\left[1-\left(\Phi+\frac{1}{\mathcal{H}}\Psi'\right)+\frac{1}{3\mathcal{H}}\partial_i V^i\right]. \tag{6.52}$$

This corrects upon the local expansion used in Ref.[198] by terms involving the potentials, in addition to the more intuitive velocity divergence. The potential terms are necessary corrections on (super-)horizon scales, which guarantees that at all times the local Hubble flow is exactly $v_{\rm ph} = H_F d_{\rm ph}$, where $v_{\rm ph}$ is the physical relative velocity and $d_{\rm ph}$ is the proper separation.

The local scale factor $a_F(\tau_F)$ may be normalized according to Eq. (6.31)

$$\frac{a_F(\tau_F)}{a(\tau_F)} \to 1 \qquad \text{as} \quad \tau_F \to 0\,, \tag{6.53}$$

for which we need the ratio of scale factors at early times at a fixed *spacetime* point

$$\lim_{\tau_F\to 0}\frac{a_F(\tau_F)}{a(\tau(\tau_F))} = 1+\frac{2}{3}\Phi_{\rm ini}\,, \tag{6.54}$$

where $\Phi_{\rm ini}$ is the initial perturbation on superhorizon scales. The local scale factor can then be obtained by direct integration

$$\frac{a_F(\tau_F)}{a(\tau)} = 1 + \frac{2}{3}\Phi_{\rm ini} + \int_0^\tau d\tau' \left(-\Psi' + \frac{1}{3}\partial_i V^i \right) . \tag{6.55}$$

The acceleration of expansion can be furthermore computed using

$$\frac{1}{a_F^2(\tau_F)} \frac{d\mathcal{H}_F(\tau_F)}{d\tau_F} = \frac{1}{3} \frac{\mathcal{H}_F(\tau_F)}{a_F(\tau_F)} \nabla_\mu U^\mu + \frac{1}{3} (e_0)^\nu \, \nabla_\nu \nabla_\mu U^\mu \, . \tag{6.56}$$

6.3.1 CFC metric

Using the general results Eqs. (6.17)–(6.19), the tidal metric correction for a long-wavelength scalar perturbation is given by

$$h^F_{00} = -\left(\partial_k \partial_l \Phi - \frac{1}{3} \delta_{kl} \partial^2 \Phi \right) x^k_F x^l_F \tag{6.57}$$

$$h^F_{0i} = \frac{2}{3} \left(\mathcal{H}' - \mathcal{H}^2 \right) \left(\delta_{kl} V_i - \delta_{il} V_k \right) x^k_F x^l_F + \frac{2}{3} \left(\delta_{ki} \partial_l - \delta_{kl} \partial_i \right) \left(\Psi' + \mathcal{H} \Phi \right) x^k_F x^l_F \tag{6.58}$$

$$\begin{aligned} h^F_{ij} = {} & \frac{1}{3} \left(\delta_{lj} \partial_i \partial_k \Psi + \delta_{li} \partial_j \partial_k \Psi - \delta_{ij} \partial_k \partial_l \Psi - \delta_{kl} \partial_i \partial_j \Psi \right) x^k_F x^l_F \\ & + \frac{2}{9} \mathcal{H} \, \partial^2 V \left(\delta_{ij} \delta_{kl} - \delta_{ik} \delta_{jl} \right) x^k_F x^l_F \, . \end{aligned} \tag{6.59}$$

First let us restrict to a spherically symmetric configuration around the CFC origin. This amounts to setting V^i, $\partial_i \Phi$ and $\partial_i \Psi$ to zero, and replacing $\partial_i \partial_j \Phi$ with $(1/3) \delta_{ij} \partial^2 \Phi$, and the

same for Ψ. The components h^F_{00} and h^F_{0i} then vanish, and h^F_{ij} is reduced to

$$h^F_{ij} = \frac{1}{3} K_F \left(x_{F,i} x_{F,j} - \delta_{ij}\, r_F^2 \right) , \tag{6.60}$$

where we have suggestively introduced *the effective spatial curvature*

$$K_F \equiv \frac{2}{3} \left(\partial^2 \Psi - \mathcal{H} \partial_i V^i \right) . \tag{6.61}$$

Given the residual gauge freedom discussed in Appendix E, we can rescale the radial coordinate

$$r_F \longrightarrow r_F \left(1 - \frac{1}{12} K_F\, r_F^2 \right) , \qquad dr_F \longrightarrow \left(1 - \frac{1}{4} K_F\, r_F^2 \right) dr_F \, . \tag{6.62}$$

In terms of this new radial coordinate, the area radius is given by $(1 + K_F\, r_F^2/4)^{-1} r_F$, and the CFC metric up to $\mathcal{O}[(x^i_F)^2]$ reads

$$ds^2 = a_F^2(\tau_F) \left[-d\tau_F^2 + \frac{\delta_{ij} dx^i_F dx^j_F}{\left(1 + K_F\, r_F^2/4 \right)^2} \right] . \tag{6.63}$$

This is exactly an FLRW spacetime with *nonzero* spatial curvature (and hence non-Euclidean geometry on each constant-scale-factor slice), parametrized in the canonical stereographic coordinates.

Next, we examine the case of more general configurations of Φ, Ψ. We will see in the next Section that for the case of the scalar metric perturbation, a necessary condition for the local spacetime to resemble a homogeneous FLRW universe is that all cosmic fluids have

the same bulk motion V_i along the central geodesic, just like the central free-falling observer. Then by way of the time-space Einstein equation, the time-space component h^F_{0i} of the tidal metric correction evaluates to zero. Without loss of generality, we shall set $h^F_{0i} = 0$ for the remainder of our discussion. We again make use of the residual gauge freedom described in Appendix E to transform h^F_{ij} into a simpler form. We consider the following choice for Eq. (E.1)

$$\begin{aligned} A_{i,jkl} &= a\left(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right)\partial^2\Psi + b\left(\delta_{ij}\partial_k\partial_l + \delta_{il}\partial_k\partial_j + \delta_{ik}\partial_l\partial_j\right)\Psi \\ &\quad + c\left(\delta_{kl}\partial_i\partial_j + \delta_{jl}\partial_i\partial_k + \delta_{jk}\partial_i\partial_l\right)\Psi, \end{aligned} \tag{6.64}$$

where the coefficients are given by $a = -1/9$, $b = 2/3$, $c = -1/3$. We furthermore perform a rescaling of the radial coordinate as done in Eq. (6.62). After some algebra, a simple form for the CFC metric with the tidal correction is derived

$$\begin{aligned} ds^2 = a_F^2(\tau_F)\Big\{ &- \left[1 + \left(\partial_k\partial_l\Phi - \frac{1}{3}\delta_{kl}\partial^2\Phi\right)x^k_F x^l_F\right]d\tau_F^2 \\ &+ \left[1 - \left(\partial_k\partial_l\Psi - \frac{1}{3}\delta_{kl}\partial^2\Psi\right)x^k_F x^l_F\right]\frac{\delta_{ij}dx^i_F dx^j_F}{\left(1 + K_F\, r^2/4\right)^2}\Big\}. \end{aligned} \tag{6.65}$$

It can be seen that this metric is in the conformal Newtonian form: the temporal and spatial directions are orthogonal to each other, and the scalar correction to h^F_{00} is related to the trace of the space-space metric h^F_{ij}, if the energy-stress tensor is without any anisotropic stress so that $\Phi = \Psi$. Eq. (6.65) may be referred to as the Poisson coordinates, in which gravitational dynamics is manifestly Newtonian on scales much smaller than the Hubble scale. Therefore, in addition to modifications of the scale factor a_F and of the spatial

curvature K_F, a general configuration for the long-wavelength perturbation results in a departure from the FLRW metric in the form of an anisotropic tidal tensor

$$K_{ij}^{\Phi} \equiv \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\partial^2\right)\Phi\,, \tag{6.66}$$

and the analogous K_{ij}^{Ψ}. It is important to highlight that while Eq. (6.66) has the identical form as the Newtonian tidal tensor, it involves the potential perturbation in the cN gauge, and remains valid on arbitrarily large scales, as long as the metric perturbation is small in amplitude.

6.3.2 Conditions for separate universe

In a homogeneous FLRW universe, the dynamics are governed by the Friedmann equations. Therefore, a CFC observer naturally expects to be able to describe the local dynamics using the Friedmann equations (note that at linear order the anisotropic tidal correction does not affect the evolution of the density), as long as she refers to locally measured quantities. We now examine under what conditions this picture precisely holds.

For practical purposes, we assume that the background FLRW universe has flat spatial geometry. According to the first Friedmann equation, the square of the Hubble rate is related to the density. In the local CFC, the long-wavelength mode modifies both the expansion rate $H \to H_F$ and the "background" density $\bar{\rho} \to \rho_F$. Inserting Eq. (6.52), we find that the local version of the Friedmann equation holds only if an extra term is supplied

with

$$H_F^2 = \frac{8\pi G}{3}\rho_F - \frac{K_F^{\rm Fr}}{a_F^2}\,. \tag{6.67}$$

Here we define

$$K_F^{\rm Fr} = \mathcal{H}^2\Delta_{\rm cN} + 2\mathcal{H}\left(\Psi' + \mathcal{H}\Phi\right) - \frac{2}{3}\mathcal{H}\,\partial_i V^i\,, \tag{6.68}$$

which behaves as the spatial curvature. Here $\Delta_{\rm cN}$ is the local value of the long-wavelength density perturbation. By way of the 00-Einstein equation for perturbations in the cN gauge, it can be verified that Eq. (6.68) represents the same quantity as Eq. (6.61). Therefore, the effective spatial curvature, depicting the local spatial geometry, also plays the expected role in the FLRW dynamics. Furthermore, in FLRW cosmology, the spatial curvature arises as a constant of integration, and therefore it is insightful to check if $K_F = K_F^{\rm Fr}$ is indeed a constant in time.

Let us first stay as general as possible. Assume that the Universe is filled with several, uncoupled (ideal) fluids, labeled by $I = 1, 2, \cdots$. Each fluid has a homogeneous equation of state $\bar{\mathcal{P}}_I = w_I\bar{\rho}_I$, with a possibly time-dependent w_I. The homogenous part of each fluid evolves as $\bar{\rho}_I' = -3\mathcal{H}(1+w_I)\bar{\rho}_I$, which translates into

$$\left(\mathcal{H}^2\Omega_I\right)' = -\left(1+3w_I\right)\mathcal{H}^3\Omega_I, \tag{6.69}$$

where $\Omega_I = \bar{\rho}_I/\bar{\rho}$ and $\bar{\rho}$ is the total (homogeneous) energy density. Furthermore, the

acceleration is given by

$$\mathcal{H}' = -\frac{1}{2}\mathcal{H}^2\left(\sum_I (1+3w_I)\Omega_I\right). \tag{6.70}$$

In the global cN-gauge coordinates, inhomogeneity of each component is described by a fractional density perturbation Δ_I. Each component also has a peculiar velocity $V_I^i = \partial^i V_I$, which is to be distinguished from the velocity of the CFC observer V^i. All fluids are sources in the 00- and $0i$-Einstein equations,

$$\partial^2\Psi = \frac{3}{2}\mathcal{H}^2\left(\sum_I \Omega_I\Delta_I\right) - \frac{9}{2}\mathcal{H}^3\left(\sum_I \Omega_I\,(1+w_I)\,V_I\right), \tag{6.71}$$

$$\Psi' + \mathcal{H}\Phi = -\frac{3}{2}\mathcal{H}^2\left(\sum_I \Omega_I\,(1+w_I)\,V_I\right). \tag{6.72}$$

Each component may have a pressure perturbation $\delta\mathcal{P}_I$ in addition to the density perturbation $\delta\rho_I = \rho_I\Delta_I$. Those two are related in the fluid's rest frame through

$$\delta\mathcal{P}_I \equiv c_{s,I}^2\,\delta\rho_I, \tag{6.73}$$

where the rest-frame speed of sound $c_{s,I}^2$ does not necessarily have to be equal to the adiabatic sound speed

$$c_{\mathrm{ad},I}^2 \equiv \frac{\bar{\mathcal{P}}_I'}{\bar{\rho}_I'} = w_I - \frac{w_I'}{3\mathcal{H}(1+w_I)}. \tag{6.74}$$

This means that we take into account the possibility of non-adiabatic pressure. Each fluid

satisfies the continuity equation and the Euler equation[212]

$$\begin{aligned}\Delta_I' + (1+w_I)\,\partial^2 V_I - 3\,(1+w_I)\,\Psi' &= 3\mathcal{H}\left(w_I - c_{s,I}^2\right)\left[\Delta_I - 3\mathcal{H}(1+w_I)V_I\right] \\ &\quad +3\mathcal{H}w_I' V_I \qquad (6.75)\\ V_I' + \left(1-3c_{s,I}^2\right)\mathcal{H}V_I + \frac{c_{s,I}^2}{1+w_I}\Delta_I + \Phi &= 0\,. \qquad (6.76)\end{aligned}$$

Using Eqs. (6.69)–(6.70), Eqs. (6.71)–(6.72), and Eqs. (6.75)–(6.76), the rate of change in Eq. (6.68) can be directly computed. For the CFC observer, V^i satisfies the equation of geodesic motion, which can be obtained by setting $c_{s,I}^2 = 0$ in Eq. (6.76)). After some algebra, we finally obtain

$$\begin{aligned}\frac{dK_F}{d\tau} &= \left[\mathcal{H}^2\left(\sum_I \Omega_I \Delta_I\right) - 3\mathcal{H}^3\left(\sum_I \Omega_I (1+w_I)V_I\right) - \frac{2}{3}\mathcal{H}\,\partial_i V^i\right]' \\ &= -\mathcal{H}^2\left[\sum_I (1+w_I)\,\Omega_I \partial_i\left(V_I^i - V^i\right)\right] + \frac{2}{3}\mathcal{H}\,\partial^2(\Phi - \Psi)\,. \qquad (6.77)\end{aligned}$$

It is straightforward to interpret when the right hand side vanishes. First, the second term vanishes if the long-wavelength perturbation does not support anisotropic stress, so that $\Phi = \Psi$, as true in FLRW cosmology. Secondly, pressureless matter ($w = c_s^2 = 0$; such as cold dark matter) always has the same geodesic bulk motion $V_I^i = V^i$ as the local observer, and hence does not contribute to the summation. Further, dark energy fluid $w_I = -1$ also does not contribute to the summation. This means that for ΛCDM cosmology, the separate universe picture will be exact along the central geodesic. If there are additional fluids of other types, then in general only if all fluids have the same geodesic motion as the CFC observer, i.e. $V_I^i = V^i$ for all I's, then the dynamics is FLRW-like.

The condition for separate FLRW evolution is approximately met if the sound horizons of all fluids are much smaller compared to the wavelength of interest, i.e. $\int a(\tau)c_s(\tau)d\tau \ll a/k_L$.[213] In that case, non-gravitational forces are negligible on the scale of interest, and geodesic bulk motion is a very good approximation. On the other hand, quintessence-like fluids with w_I very close to -1 would also have negligible contribution to the summation, and is not expected to severely violate the separate universe condition. Those have been theorized to explain the Dark Energy that induces the recent cosmic acceleration.

Now assuming that all cosmic fluids co-move with the free-falling observer, such that $V_i = \delta T^0{}_i/(\bar{\rho} + \bar{\mathcal{P}})$, we can re-write Eq. (6.68) using Eqs. (6.71)–(6.72) and find

$$K_F = \frac{2}{3}\left(\partial^2\Psi - \frac{\mathcal{H}\,\partial^i\delta T^0{}_i}{\bar{\rho}+\bar{\mathcal{P}}}\right) = \frac{2}{3}\partial^2\mathcal{R}, \tag{6.78}$$

where $\mathcal{R}$ is the gauge-invariant *curvature perturbation on comoving slices*.[129] The conservation of K_F is hence directly related to the conservation of $\mathcal{R}$ at linear order in perturbation. Note that the condition for a constant $\mathcal{R}$ differs from that for a constant ζ, the curvature perturbation on *uniform density slices*. It is found that the latter is conserved if pressure is only a function of energy density.[214] The conservation of K_F, however, does not exclude non-adiabatic pressure.

Finally, in Appendix F we show that the local version of the second Friedmann equation also holds given the result for H_F. Unlike the first Friedmann equation, no additional condition is imposed on the matter content of the Universe.

6.4 Squeezed-limit matter bispectrum

We now study how the evolution of small-scale density perturbation is influenced by a background density perturbation of very long wavelength. As we have argued at the beginning of this Chapter, the small-scale perturbation effectively evolves in an FLRW universe, with modified expansion history and an effective spatial curvature, plus an anisotropic tidal force. We therefore first restore small-scale scalar perturbations in the CFC metric we have derived for the coarse-grained universe. It is most convenient to work within the Poisson coordinates

$$ds^2 = a_F^2(\tau_F)\Bigg\{ -\left[1+2\phi+\left(\partial_k\partial_l\Phi-\frac{1}{3}\delta_{kl}\partial^2\Phi\right)x^k x^l\right]d\tau_F^2 + \left[1-2\psi-\left(\partial_k\partial_l\Psi-\frac{1}{3}\delta_{kl}\partial^2\Psi\right)x^k x^l\right]\frac{\delta_{ij}dx^i dx^j}{(1+K_F\, r^2/4)^2}\Bigg\}, \tag{6.79}$$

and the small-scale modes are chosen to be parametrized by two scalar gravitational potentials ϕ and ψ in the cN gauge [2]

We now want to solve for the evolution of ϕ, ψ in the presence of the long-wavelength mode Φ, Ψ. In the traditional second-order perturbation theory, one would have to solve the nonlinear Einstein equations. However, according to the discussion in Section 6.2.2, the CFC formalism is advantageous that this is in fact not necessary when we restrict to the regime where short-scale perturbations have long entered the horizon. The wavelength of the large-scale mode, on the other hand, is not limited to subhorizon scales. This regime,

[2] Note that here we refer to the gauge choice for short-wavelength perturbations. It should not be confused with the gauge choice for long-wavelength perturbations, which is by construction gauge fixed in the CFC.

though not entirely general, is actually of the most practical interest. The reason is that, with adiabatic initial perturbations, no interesting dynamics take place at all until the short-wavelength modes enter the horizon. The long-wavelength mode, still far outside the horizon at that moment, do not have any dynamical impact.

Modification of the background quantities a_F and K_F by the long mode can be simply accounted for in the background part of the Einstein tensor. Moreover, for subhorizon $k_S \gg \mathcal{H}$, it is consistent to set $\psi = \phi$ even at nonlinear orders. What is more, only *scalar* Einstein equations, including the 00-component as well as the divergence of the $0i$-component, will be used to solve the evolution of the matter density. Finally, it is sufficient to solve the equations right on the central geodesic G, since we are free to choose any desired fluid trajectory to be the CFC central geodesic, provided that during the analytical manipulation no further spatial derivative is taken. By construction, only double spatial derivatives of the long-wavelength perturbation Φ and Ψ will enter the dynamical equations in CFC. At second order in perturbation, the only terms in the Einstein equations have to be of the form

$$F(\phi)\left(\partial_i\partial_j\Psi - \frac{1}{3}\delta_{ij}\partial^2\Psi\right), \tag{6.80}$$

and correspondingly for ψ and Ψ. To contribute to the scalar equations, any tensorial structure must contract with δ_{ij}, which then vanishes due to the trace-free structure. We therefore conclude that in the CFC no explicit second-order long-short coupling terms should appear in the Einstein equations.

6.4.1 The calculation in CFC

Provided that we are focusing on $k \gg \mathcal{H}$, the complete set of equations describing the clustering of small-scale density modes in the background of a long-wavelength mode may be chosen to include the 00-Einstein equation, the continuity equation and the Euler equation. They read

$$\begin{aligned} \partial^2\phi &= \frac{3}{2}\mathcal{H}_F^2\,\Omega_m^F\,\delta \\ \delta' + \partial_i[v^i + (1+\delta)v_L^i] &= 0 \\ v_i' + \mathcal{H}_F\,v_i + (v_L\cdot\partial)v_i + (v\cdot\partial)v_{L,i} &= -\partial_i\phi\,, \end{aligned} \tag{6.81}$$

where δ and v_i are the small-scale density field and peculiar velocity, respectively. To simplify the notation, we have used x^i for the CFC x_F^i. We have also splitted the velocity into a small-scale component v^i and a large-scale component v_L^i. The latter, linearly induced by the long mode, satisfies the following geodesic equation

$$v_{L,i}{}' + \mathcal{H}_F\,v_{L,i} + K_{ij}^{\Phi}\,x^j = 0\,, \tag{6.82}$$

where the tidal force follows from the previous result Eq. (6.79)

$$K_{ij}^{\Phi} = \left[\partial_i\partial_j - \frac{1}{3}\delta_{ij}\partial^2\right]\Phi(\mathbf{0},\tau_F)\,, \tag{6.83}$$

The solution of Eq. (6.82) can be found by direct integration

$$v_{L,i}(\tau) = -\frac{x^j}{a_F(\tau)} \int_0^\tau a_F(\tau') K^\Phi_{ij}(\tau') d\tau' = -\frac{2x^j}{3\mathcal{H}_F} \left(\partial_i \partial_j - \frac{1}{3}\delta_{ij}\partial^2 \right) \Phi_{\rm ini}. \tag{6.84}$$

For an EdS background, the matter density is modified locally in the CFC so that

$$\Omega^F_m - 1 = \frac{K_F}{a_F^2 H_F^2} = \frac{K_F}{\mathcal{H}_F^2}\,, \tag{6.85}$$

where the effective curvature K_F is given in Eq. (6.66). Note that the departure from Euclidean geometry due to K_F is negligible $\sim H^2/K_F \sim (aH/k)^2\Phi$ on subhorizon scales with regard to structure formation, and the major effect of K_F is to modify the evolution of ρ_F, and thus affect the rate of clustering.

We have explicitly seen in Eq. (6.81) that the equations are linearized in the gravity part (the linearized Einstein tensor and the Newtonian force), but are in general nonlinear in fluid variables δ and v^i. The long mode enters through modified FLRW parameters $\mathcal{H}_F$ and Ω^F_m, as well as the tidal force in the Euler equation. Note that this is a non-trivial advantage of the CFC frame, because we have not restricted the wavelength of the long mode to be much shorter than the Hubble scale, so that Newtonian equations in principle do not hold. In other words, the CFC formalism allows us to solve the evolution of small-scale physics in the presence of background modes with arbitrarily large wavelength, but still to be able to use equations in the Newtonian form. Also note that we are seeking a solution of Eq. (6.81) right on the central geodesic, so that we stop at zeroth-order in x^i.

The only exception is the Euler equation, for which we keep terms linear in x^i as well, since solving the density requires an equation for the velocity convergence $\theta = \partial_i v^i$, which follows from the divergence of the Euler equation.

Eq. (6.81) will be solved perturbatively. We split the short-scale variables, δ, v^i and ϕ, into a linear part, corresponding to the solution in the absence of the long mode, and a second-order correction, due to its coupling with the long mode. For instance, we write

$$\delta = \delta^{(1)} + \delta^{(2)}, \tag{6.86}$$

where $\delta^{(1)}$ is the linear solution and $\delta^{(2)}$ is the second-order correction. We have introduced the proper velocity divergence $\hat{\theta} \equiv a_F^{-1}\partial_i v^i$, which involves the local scale factor. It is favorable to use the CFC proper time t_F (derivative denoted as a dot) instead of τ_F as the time variable, because we would like to compare δ with $\delta^{(1)}$ at fixed proper time.

In an EdS universe with adiabatic initial condition, the potential is constant $\phi^{(1)} = \phi_{\rm ini}$. The linear solutions on subhorizon scales are well known

$$\delta^{(1)} = \frac{2}{3\mathcal{H}^2}\partial^2\phi_{\rm ini}, \quad v_i^{(1)} = -\frac{2}{3\mathcal{H}}\partial_i\phi_{\rm ini}, \quad \hat{\theta}^{(1)} = -\frac{2}{3a\mathcal{H}}\partial^2\phi_{\rm ini}\,, \tag{6.87}$$

which are derived for the global expansion rate $\mathcal{H}$ and zero curvature $\Omega_m = 1$. We then separate the second-order part of the continuity equation and Euler equation, which encodes

the long-short mode coupling

$$\begin{aligned}
\dot{\delta}^{(2)} + \hat{\theta}^{(2)} &= -a_F^{-1} v_L^i \, \partial_i \delta^{(1)} \\
\dot{\hat{\theta}}^{(2)} + 2H\hat{\theta}^{(2)} + \frac{3}{2}H^2\delta^{(2)} &= -a_F^{-1} v_L^i \, \partial_i \hat{\theta}^{(1)} - 2a_F^{-2}(\partial_i v_L^j)(\partial_j v_{(1)}^i) \\
&\quad -2\left(H_F - H\right)\hat{\theta}^{(1)} - \frac{3}{2}\left(\Omega_m^F H_F^2 - H^2\right)\delta^{(1)} \,, \qquad (6.88)
\end{aligned}$$

Note that on the right hand sides source terms are quadratic in linear solutions Eq. (6.87). In keeping with our analysis of the long-short coupling, we only have to retain quadratic terms that are products of the short-scale mode and the long-wavelength mode. What is more, the coefficient functions must be evaluated at fixed proper time t_F; this does not imply fixed conformal time τ_F, as it depends on the choice of a_F. For example, $H_F(t_F)$ is the local expansion rate at fixed t_F, which differs from the background expansion rate $H(t_F) = 2/(3t_F)$.

At initial time we set the second-order correction to be zero $\delta^{(2)}|_{\rm ini} = \hat{\theta}^{(2)}|_{\rm ini} = 0$, and the same for their time derivatives. This means that there is no *physical* coupling between the long mode and the short mode in the CFC frame, a generic prediction of single-clock inflation.[174] By "physical coupling" we mean that small-scale perturbations at fixed physical scale is not being modulated by the long-wavelength mode. By contrast, in global coordinates we must set up the coupling between the long mode and the short mode in accord with the inflationary consistency relation,[20,161] which merely reflects the fact that a long-wavelength perturbation also modulates proper distance scales. In this sense, the CFC formalism is advantageous because it is free from general relativistic gauge artifacts

that do not represent any physical coupling.

We now turn to the computation of the second-order corrections. Since we compute to linear order in the long mode, we may compute separately the isotropic effect and the purely anisotropic effect.

Isotropic effect: local expansion and curvature

An isotropic configuration of the long mode does not induce any tidal force $K^{\Phi}_{ij} = 0$, so that $v^i_L = 0$. Ignoring those terms in Eq. (6.88), we then combine the time derivative of the continuity equation and the Euler equation to find a second-order equation for $\delta^{(2)}$,

$$\begin{aligned} \ddot{\delta}^{(2)} + 2H(t_F)\dot{\delta}^{(2)} - \frac{3}{2}H^2(t_F)\delta^{(2)} &= -2\left[H_F(t_F) - H(t_F)\right]\dot{\delta}^{(1)} \\ &\quad + \frac{3}{2}\left[H_F^2(t_F)\Omega_m^F(t_F) - H^2(t_F)\right]\delta^{(1)}, \end{aligned} \tag{6.89}$$

It boils down to computing the coefficient functions on the right hand side. We first combine Eq. (6.52) and Eq. (6.61) to obtain

$$H_F(t_F) - H(t) = -H\left(\Phi + \frac{2}{9\mathcal{H}^2}\partial^2\Phi\right), \tag{6.90}$$

$$H_F^2(t_F)\,\Omega_m^F(t_F) - H^2(t) = \frac{K_F}{a_F^2} + H_F^2(t_F) - H^2(t) = -2H^2\Phi + \frac{2}{3a^2}\partial^2\Phi. \tag{6.91}$$

However, here the CFC quantities and the global quantities are evaluated at proper time t_F and global time t, respectively. The two times correspond to the same spacetime point on the central geodesic, one in the CFC frame and the other in the global frame, but in

general differ numerically. We thus need to convert to the same numerical value. We derive for an EdS universe $H(t) - H(t_F) = -\dot{H}\,\Delta t$, where $\Delta t = t_F - t = 2\Phi/3H$. Using this result, Eqs. (6.90)–(6.91) can be converted into

$$H_F(t_F) - H(t_F) = -\frac{2H}{9\mathcal{H}^2}\partial^2\Phi, \tag{6.92}$$

$$H_F^2(t_F)\,\Omega_m^F(t_F) - H^2(t_F) = \frac{2}{3a^2}\partial^2\Phi. \tag{6.93}$$

Inserting these into Eq. (6.89), we find

$$\ddot{\delta}^{(2)} + 2H\dot{\delta}^{(2)} - \frac{3}{2}H^2\delta^{(2)} = \left(\frac{13}{9}\frac{1}{a^2}\partial^2\Phi\right)\delta^{(1)}\,. \tag{6.94}$$

The solution reads

$$\delta^{(2)} = \left(\frac{26}{63}\frac{1}{\mathcal{H}^2}\partial^2\Phi\right)\delta^{(1)} = \frac{13}{21}\,\Delta_{\rm sc}\,\delta^{(1)}\,, \tag{6.95}$$

where we have expressed the final result in terms of the long-wavelength density in the synchronous-comoving (sc) gauge $\Delta_{\rm sc} = (2/3\mathcal{H}^2)\partial^2\Phi$. This result agrees perfectly with the result from the Newtonian spherical collapse of an overdense, uniform top-hat as a curved FLRW universe,[198,215] provided that we identify the top-hat overdensity with the long-wavelength density perturbation in the sc gauge so that it is also valid if the long wavelength is not much shorter than the Hubble distance.

Anisotropic effect: tidal force

The purely anisotropic component of the local configuration of the long-wavelength perturbation does not modify the local expansion rate and does not induce a spatial curvature. To isolate the anisotropic effect, therefore, we set $H_F \to H$, $a_F \to a$, and $\Omega_m^F \to \Omega_m = 1$. On the other hand, we should keep v_L^i, the long-wavelength velocity component induced by the anisotropic tidal force in the CFC. Then Eq. (6.88) reduces to the equations[174]

$$\begin{aligned} \dot{\delta}^{(2)} + \hat{\theta}^{(2)} &= -\,a^{-1}\, v_L^i\, \partial_i \delta^{(1)} \\ \dot{\hat{\theta}}^{(2)} + 2H\,\hat{\theta}^{(2)} + \frac{3}{2}H^2\delta^{(2)} &= -\,a^{-1} v_L^i\, \partial_i \hat{\theta}^{(1)} - 2a^{-2}\left(\partial^i v_L^j\right)\left(\partial_j v_i^{(1)}\right). \end{aligned} \tag{6.96}$$

Inserting the expressions for linear perturbations Eq. (6.87) and Eq. (6.84), but then only retaining terms non-vanishing at $x^i = 0$, we find the correction to the density

$$\delta^{(2)} \;=\; \frac{4}{7}\, K_{ij}^{\Delta}\left(\frac{\partial^i\partial^j}{\partial^2}\delta^{(1)}\right), \tag{6.97}$$

where we define the anisotropic density tensor

$$K_{ij}^{\Delta} \;=\; \frac{2}{3\mathcal{H}^2}K_{ij}^{\Phi} = \left(\frac{\partial_i\partial_j}{\partial^2} - \frac{1}{3}\delta_{ij}\right)\Delta_{\rm sc}\,. \tag{6.98}$$

6.4.2 Eulerian density

We can now combine the result of Eq. (6.95) and Eq. (6.97) to find the complete density field at second-order

$$\delta_G \;=\; \left(1+\frac{13}{21}\Delta_{\rm sc}\right)\delta^{(1)}+\frac{4}{7}\,K^{\Delta}_{ij}\,\left(\frac{\partial^i\partial^j}{\partial^2}\delta^{(1)}\right). \tag{6.99}$$

The label $_G$ signifies that the result is only valid right on the central geodesic G. Indeed, the calculation within the CFC frame significantly simplifies if we only seek solutions right on G, i.e. always at the spatial origin but at all times. While the result at $x^i = 0$ is simple, it is far from sufficient. After all, we are interested in the density field not only at a single point. To derive the correlation functions for the small-scale density field, for instance the two-point correlation function or the power spectrum in Fourier space, the local observer needs the knowledge of the density field at least across a finite region around her, i.e. at finite x^i.

To address this problem, we could compute the CFC metric to higher orders in x^i, and derive the correct dynamical equations also to higher orders in x^i, and then solve for the perturbation variables, again to higher orders in x^i. Although conceptually this is straightforward, in practice this would be extremely cumbersome. The alternative is to realize that we are free to choose the trajectory of any fluid element of interest to be the central geodesic and then construct an appropriate CFC around it. That is to say, for any point Q next to the spatial origin P of a given CFC, it is possible to choose a different geodesic G' that passes through Q, and then compute the density at Q within the new CFC

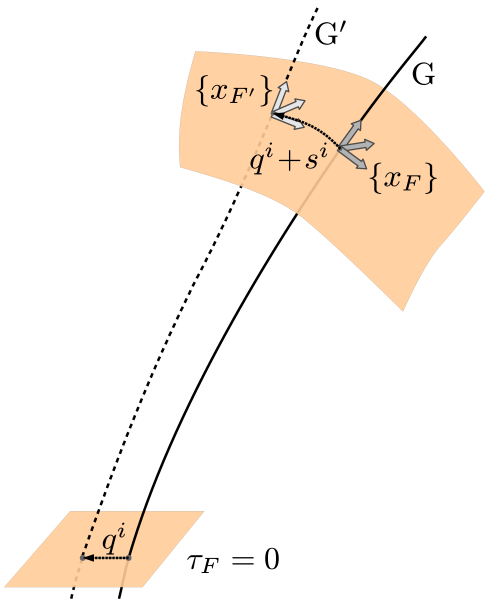


Figure 6.2: Illustration of the mapping between nearby CFC frames $\{x_F\}$, $\{x_{F'}\}$ (tetrads indicated as arrows) constructed for the geodesics G, G', respectively (thick solid and dashed). The initial separation at $\tau_F = 0$ of the two geodesics in the $\{x_F\}$ frame is given by $\mathbf{q}$ (dotted arrow), which at late times is modified to $\mathbf{q} + \mathbf{s}$ due to the local geodesic deviation.

frame constructed around G'. This is illustrated in Figure 6.2.

The remaining question is how we can translate the result for density field at Q in the new CFC frame based G' to the result at density field at Q in the origin CFC frame based on G (may be called *the central CFC frame*), so that in G, we can obtain a result for the density field across a finite region surrounding $x^i = 0$. Note that a result for a perturbation field on the central geodesic, such as Eq. (6.99), is eventually expressed in terms of the initial perturbation fields $\Phi_{\rm ini}(\mathbf{x})$ and $\phi_{\rm ini}(\mathbf{x})$ generated through the inflationary mechanism. Those are functions of the comoving position $\mathbf{x}$ at early times (before perturbations have entered the horizon and become dynamical), which we may alternatively refer to as *the Lagrangian coordinate* $\mathbf{q}$. Without loss of generality, we set the Lagrangian coordinate at early times on G to be zero, and then the Lagrangian coordinate on G' gives its initial displacement from

G. At later times, the separation between the two geodesics will vary $\mathbf{x} = \mathbf{q} + \mathbf{s}(\mathbf{q})$ because of geodesic deviation. It boils down to the calculation of the shift vector $\mathbf{s}$ as a function of the Lagrangian position $\mathbf{q}$.

Note that we have $\mathbf{q} = 0$ in the absence of long-wavelength perturbations and their induced tidal correction $h^F_{\mu\nu}$. This is because in a homogeneous FLRW universe a comoving separation is conserved. Therefore, it would be convenient to write down the geodesic deviation with respect to the conformal metric $\tilde{g}_{\mu\nu}$ (see Eq. (6.3)) instead of the physical metric. Then the observer's trajectory would not be a "geodesic" of the conformal metric $\tilde{g}_{\mu\nu}$. Let us use a convenient affined paramter $d\tilde{\lambda} = a_F^{-2} dt_F$. A physical geodesic equation for G can be rewritten with respect to the conformal metric

$$\frac{d^2x^\mu}{d\tilde{\lambda}^2} + \tilde{\Gamma}^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tilde{\lambda}}\frac{dx^\beta}{d\tilde{\lambda}} = \tilde{g}_{\alpha\beta}\frac{dx^\alpha}{d\tilde{\lambda}}\frac{dx^\beta}{d\tilde{\lambda}}\tilde{g}^{\mu\gamma}\partial_\gamma \ln a_F \,. \tag{6.100}$$

Because along the central geodesic $\tilde{\Gamma}^F = 0$, the equation reduces to

$$\frac{d}{d\tilde{\lambda}}\hat{e}^\mu = a_F^2 \mathcal{H}_F \delta^\mu{}_0 \,, \tag{6.101}$$

where $\hat{e}^\mu = a_F^2 (e_0)^\mu$ is a vector tangent to G (rescaled from the 4-velocity vector of the CFC observer). When the curve is not a "geodesic", the geodesic deviation equation should be modified to[211]

$$\frac{d}{d\tilde{\lambda}}\left(\frac{d}{d\tilde{\lambda}}x^\alpha\right) = \hat{e}^\beta \hat{e}^\gamma (\tilde{R}_F)^\alpha{}_{\beta\gamma\mu}\, x^\mu + x^\gamma \tilde{\nabla}_\gamma \left(\frac{d}{d\tilde{\lambda}}\hat{e}^\alpha\right) . \tag{6.102}$$

Here x^α is a vector that points from G to G' and is perpendicular to the tangent direction

of G. It quantifies the deviation of G' from G at each moment along G. This covariant equation can specialize to the CFC frame, where the metric and connection coefficients greatly simplify on the central geodesic G,

$$a_F \left(a_F x^{\alpha\prime}\right)' = a_F^2 (\tilde{R}_F)^{\alpha}{}_{00\mu}\, x^\mu + \delta^\alpha{}_0\, x^0 \left(a_F^2 \mathcal{H}_F\right)' , \quad \text{or}$$

$$a_F^{-1}(a_F x^{\alpha\prime})' - \delta^\alpha_{\ 0}\, x^0 \left(\frac{a_F''}{a_F} + \mathcal{H}_F^2\right) = (\tilde{R}_F)^{\alpha}{}_{00\mu}\, x^\mu \,. \tag{6.103}$$

With the relevant curvature components $(\tilde{R}_F)^0{}_{00\mu} = -(\tilde{R}_F)_{000\mu}$ vanishing in the CFC, the temporal component $\alpha = 0$ satisfies the equation

$$a_F^{-1}(a_F x^{0\prime})' - x^0 \left(\frac{a_F''}{a_F} + \mathcal{H}_F^2\right) = 0 \,. \tag{6.104}$$

Without explicit dependence on the long-wavelength mode in this homogeneous differential equation, it is a matter of definition to set $x^0 = 0$ at all times. Also, the value of x^0 does not enter the equation for the spatial displacement x^i, which is given by

$$a_F^{-1}(a_F x^{i\prime})' = \left(\tilde{R}_F\right)^i{}_{00j}\, x^j, \tag{6.105}$$

with initial condition $\mathbf{x}(\tau_F \to 0) = \mathbf{q}$. Then this equation can be solved perturbatively, accurate to linear order in the long mode. We find the shift vector to be

$$\begin{aligned} s^i(\tau_F) = x^i(\tau_F) - q^i &= \int_0^{\tau_F} \frac{d\tau'}{a_F(\tau')} \int_0^{\tau'} d\tau'' a_F(\tau'') (\tilde{R}_F)^i{}_{00j}\, q^j \\ &= \frac{1}{2} \int_0^{\tau_F} \frac{d\tau'}{a_F(\tau')} \int_0^{\tau'} d\tau'' a_F(\tau'') \left(\partial_i \partial_j h_{00}^F\right) q^j \,. \end{aligned} \tag{6.106}$$

In the second line, we have used previous result for the CFC tidal metric corrections Eq. (6.17). Eq. (6.106) is the correct result at first order in the long-wavelength perturbation, and is gauge invariant due to the uniqueness of the CFC frame. This establishes a relation between the Lagrangian position $\mathbf{q}$ of the neighboring geodesic G', and its position in the central CFC frame at a given later time, i.e.

$$x_F^i = q^i + s^i(\mathbf{q}, \tau_F) = q^i + \frac{1}{2}\int_0^{\tau_F} \frac{d\tau'}{a_F(\tau')} \int_0^{\tau'} d\tau'' a_F(\tau'') \left(\partial_i \partial_j h_{00}^F\right) q^j \,. \tag{6.107}$$

We can then relate physical quantities measured in the local CFC frame to that measured in the central CFC frame. The latter, of course, is what we seek to derive. For a scalar quantity such as the matter density field, the transformation is a simple shift in the assignment of spatial coordinates. We therefore have

$$\rho_F(x_F, \tau_F) = \rho_{F(\mathbf{q})}(\tau_F)\Big|_{\mathbf{q}=\mathbf{x}_F - \mathbf{s}(\mathbf{x}_F, \tau_F)} = \left[1 - s^i(\mathbf{x}_F, \tau_F)\,\partial_i\right] \rho_{F(\mathbf{q})}(\tau_F)\Big|_{\mathbf{q}=\mathbf{x}_F} \,, \tag{6.108}$$

We will call $\rho_F(x_F, \tau_F)$ the Eulerian density field in the CFC, as different from the Lagrangian result Eq. (6.99), which only holds right on the central geodesic $x_F^i = 0$. Using

We now immediately apply the above discussion to the specific case of long-wavelength scalar perturbations. In that case, Eq. (6.106) evaluates to

$$\begin{aligned} s^i &= \frac{1}{2}\, q^j \int_0^{\tau_F} \frac{d\tau'}{a_F(\tau')} \int_0^{\tau'} d\tau'' a_F(\tau'') \left(\partial_i \partial_j h_{00}^F\right) \\ &= -\frac{2}{3\mathcal{H}^2}\, K_{ij}^{\Phi}\, x^j = -K_{ij}^{\Delta}\, x^j. \end{aligned} \tag{6.109}$$

The Eulerian density field $\delta_E = \delta_G - s^i\,\partial_i\delta_G$ is then given by

$$\delta_E \;=\; \delta^{(1)} + \frac{13}{21}\Delta_{\rm sc}\,\delta^{(1)} + \frac{4}{7}\,K^{\Delta}_{ij}\,\left(\frac{\partial^i\partial^j}{\partial^2}\delta^{(1)}\right) + K^{\Delta}_{ij}\,x^i\partial^j\delta^{(1)}\,. \tag{6.110}$$

The last term, which we refer to as *the displacement term*, is the $\mathcal{O}[x^i]$ correction that generalizes the result at $x^i = 0$ to leading order in finite x^i. The presence of terms explicitly dependent on x^i is in sharp constrast with the calculation carried out in global coordinates, usually referred to as the Standard Perturbation Theory (SPT). Otherwise, the nonvanishing terms at $x^i = 0$ are very similar to the result of SPT. Nevertheless, we point out two crucial differences between the CFC result and the well known SPT result. First, the SPT result is derived in the subhorizon limit (for all length scales involved) assuming Newtonian dynamics. The CFC result presented in Eq. (6.110), however, is applicable to large-scale modes of arbitrary wavelength, as long as $\Delta_{\rm sc}$ and K^{Δ}_{ij} are in the sc gauge (the gauge choice for the small-scale modes is not important in the subhorizon regime $k \gg \mathcal{H}$), which always scale as double spatial derivatives of the primordial potential, rather than the potential itself. The second point is that the displacement term appears only as a relative displacement $\propto x^i$ from the CFC observer, who herself is also moving. The bulk velocity of the entire CFC frame enters the SPT result, but has no local effect whatsoever on the small-scale physics due to the Galilean Principle. These points highlight the advantage that in the CFC gauge artifacts are removed by construction. Note also that after a Fourier transform the displacement term has a correction to the power spectrum proportional to the spectral slope of the linear power spectrum.

Eq. (6.110) still does not have exactly the same numerical coefficients as in the SPT result. The discrepancy can be accounted for by two effects. First, in the local CFC frame, density constrast has been defined with respect to the local average density ρ_F, which at fixed proper time differs from the global average $\bar{\rho}$

$$\rho_F(t_F) \;=\; \bar{\rho}(t_F)\,[1+\Delta_{\rm sc}(t_F)]\,, \tag{6.111}$$

due to the long-wavelength density perturbation. If we choose to normalize to $\bar{\rho}$ in keeping with the SPT convention, an additional correction $\Delta_{\rm sc}\delta^{(1)}$ must be added to Eq. (6.110). The second effect has something to do with the fact that we always refer to fixed comoving scale k when evaluating the short-scale perturbation. This rescales from the physical scale by the local scale factor a_F, while in SPT the global scale factor a is used. At fixed proper time, they differ by

$$a_F(t_F) \;=\; a(t_F)\,[1-\Delta_{\rm sc}(t_F)/3]\,. \tag{6.112}$$

Converting back to the comoving scale normalized to the global scale factor is equivalent to adding back the trace part of the tensor K^{Δ}_{ij}. In Appendix G, we show that the CFC result Eq. (6.110) reproduces the SPT result, once the aforementioned two effects are corrected for.

6.4.3 Squeezed-limit matter bispectrum

Parallel to the derivation in Section 5.4.2, we can Fourier transform with respect to the short-wavelength scale in Eq. (6.110), and obtain a local power spectrum for the small-scale perturbations as modulated by a long-wavelength mode $\Delta(\mathbf{k}_L)$,

$$P_\delta(\mathbf{k}_S)|_{\Delta(\mathbf{k}_L)} = \left[1 + K^\Delta_{ij}\,\hat{k}^i_S\hat{k}^j_S\left(\frac{8}{7} - \frac{d\ln P_\delta(k_S)}{d\ln k_S}\right) + \frac{26}{21}\Delta_{\rm sc}\right] P_\delta(k_S). \quad (6.113)$$

Here $P_\delta(k_S)$ is the linear power spectrum in the absence of the long-wavelength perturbation. We see that the local power spectrum is modulated by the local value of the long-wavelength density in the sc gauge. Given that $\Delta(\mathbf{k}_L)$ is also stochastic, the same physics encapsulated by the modulated local power spectrum Eq. (6.113) may be rephrased in terms of the following *locally observable squeezed matter bispectrum*

$$\langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\Delta_{\rm sc}(\mathbf{k}_L)\rangle' = \left[\frac{26}{21} + \left(\mu^2_{SL} - \frac{1}{3}\right)\left(\frac{8}{7} - \frac{d\ln P_\delta(k_S)}{d\ln k_S}\right)\right] P^\Delta_{\rm sc}(k_L)P_\delta(k_S). \quad (6.114)$$

Here $\mathbf{k}_S = (\mathbf{k}_1 - \mathbf{k}_2)/2$ and $\mu_{SL} \equiv \hat{\mathbf{k}}_L \cdot \hat{\mathbf{k}}_S$ and we require the squeezed limit $|\mathbf{k}_1| \approx |\mathbf{k}_2| \approx k_S \gg k_L$. The prime reminds us that an overall factor $(2\pi)^3\delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_L)$ due to statistical homogeneity has been omitted.

Eq. (6.114) can be interpreted as the following. Imagine that there are (imaginary) observers distributed around the Universe across very large scales, who follow geodesic motion in a coarse-grained universe (i.e. only respond to the long-wavelength perturbation). Each of them also carries a clock that measures the observer's proper time since the begin-

ning of the Universe. Then these observers can record *at fixed proper time* the clustering power spectrum of small-scale density fluctuation *at fixed physical wavelength* in their own vicinity. The information from all of them are then communicated through spacetime to reach a master observer at the present epoch, who makes an estimate for the ensemble average $\langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\Delta_{\rm sc}(\mathbf{k}_L)\rangle$, based on the information collected. Under this interpretation, Eq. (6.114) may be regarded as a squeezed matter bispectrum defined on synchronous slices. However, note that Eq. (6.114) is applicable to a region of space more extensive than the wavelength of the long mode, even the entire Universe, despite that the CFC has limited range of validity. In deriving this result, the CFC construction is only utilized as a theoretical tool.

We comment that Eq. (6.114) still does not correspond to an experimental observable in real LSS surveys. In a perturbed universe, observers are ignorant of the true physical position of LSS tracers. Rather, the position is inferred from the observed redshift and apparent direction in the sky, which differs from the physical position because photon propagation from the source objects is distorted by metric perturbations according to general relativity. This has been discussed in Section 5.4.1 as the projection effects. We stress that both the locally observed clustering statistics Eq. (6.114) and further correction to LSS observable due to projection are well-defined and are separately gauge invariant.

6.4.4 Scale-dependent bias

Tracers of matter distribution in the Universe, such as galaxies, typically form out of density peaks,[216] and their abundance is sensitive to the clustering power spectrum of matter

perturbations on scales relevant to their formation (typically $k > 0.1\,h\,\mathrm{Mpc}^{-1}$) prior to their formation. If the small-scale power spectrum is modulated by a long-wavelength density perturbation due to nonlinear evolution, the large-scale tracer abundance Δ_g is expected to correlate with the long-wavelength mode Δ, typically in the form of a clustering bias $\Delta_g = b_g\Delta$. This affects the clustering of tracers on large scales.

It might be that small-scale power spectrum is modulated by long-wavelength perturbations already in the initial condition, as suggested by a variety of early universe scenarios. Regardless of the details of the physics, it is conventional to parametrize the non-Gaussianity in the primordial potential perturbation[88]

$$\Phi_{\rm NG}\Big|_{\rm ini} \;=\; \Phi + f_{\rm NL}^{\rm loc}\left(\Phi^2 - \left\langle\Phi^2\right\rangle\right), \tag{6.115}$$

where Φ has Gaussian statistics, and $f_{\rm NL}^{\rm loc}$ quantifies the non-Gaussian statistics of the physical potential field $\Phi_{\rm NG}$. This leads to a density field that is non-Gaussian but grows linearly

$$\delta \;=\; \left[1 + 2f_{\rm NL}^{\rm loc}\,\Phi\right]\,\delta^{(1)}, \tag{6.116}$$

where $\delta^{(1)}$ is gaussian. We see that the small-scale fluctuation is modulated by a large-scale gravitationl potential.

Such modulation leads to a contribution to the tracer clustering bias on large scales for

EdS cosmology[22]

$$\Delta b(k_L) \;=\; 2(b_g-1)f_{\rm NL}^{\rm loc}\delta_c\,\frac{3\mathcal{H}^2}{2k_L^2}, \tag{6.117}$$

where δ_c is some threshold for collapse and (b_g-1) is the *Lagrangian* bias in the absence of primordial non-Gaussianity. This is a scale-dependent bias as it is enhanced on large-scales $\Delta b(k_L)\propto 1/k_L^2$, a scaling inherited from the fact that in Eq. (6.116) the modulation is proportional to the large-scale potential Φ, rather than the large-scale density contrast (in the sc gauge) $\Delta_{\rm sc}\sim(k_L^2/\mathcal{H}^2)\Phi$. Therefore, primordial non-Gaussianity of the type Eq. (6.115) leaves a prominent signature in the clustering of LSS tracers as it greatly enhances the clustering power on large scales.

It is predicted that in standard single-field inflation, there is no physical modulation of the type Eq. (6.115) in the initial condition. Further, our result Eq. (6.113) shows that late-time nonlinear gravitational clustering only induces a modulation by the (sc gauge) large-scale density $\Delta_{\rm sc}$ rather than the potential itself, which then only leads to a scale-independent clustering bias. Therefore, we come to the important conclusion that *in standard single-field inflation, there is no scale-dependent clustering bias on large scales arising from general relativistic clustering.*[162,163] This has important implication for experiments in an exciting era when the next generation of LSS surveys are designed to target at $f_{\rm NL}^{\rm loc}\sim\mathcal{O}(1)$ to test various models of inflation.[217] This is expected from *the Equivalence Principle* in general relativity — a homogeneous metric perturbation can be eliminated by a change of spacetime coordinates, and only the spatial variation of the metric perturbation has an

impact on local physical processes. Our conclusion is in disagreement with a series of previous statements in the literature that a scale-dependent bias corresponding to an effective $f_{\rm NL}^{\rm loc} = -5/3$ is predicted by general relativity even for single-field inflation.[218–222] Our result also implies that the detection of a scale-dependent clustering bias $\propto 1/k_L^2$ on large scales will be compelling evidence for non-Gaussianity in the initial condition of the Universe, invalidating single-clock inflation.

At last, we put caveat that photon projection effects can induce apparent clustering of tracers on large scales. Since photon propagation does not occur locally, but takes place across cosmological distances that are comparable or larger than the wavelength of long-wavelength perturbations, this apparent clustering has a component that does scale like an effective $f_{\rm NL}^{\rm loc}$ parameter.[162] However, this apparent clustering is purely geometrical, and therefore is not enhanced by the clustering bias of the type of tracer of interest. In other words, while genuine primordial non-Gaussianity will preferentially enhance the abundance of the most biased tracer species, projection effects induce the same amount of clustering for all types of tracers[168 3]. In principle, the apparent clustering from projection effects can be distinguished from the physical tclustering induced by primordial non-Gaussianity.

6.5 Summary

In this Chapter, we have systematically studied how the dynamics of small-scale perturbations can be locally affected by a long-wavelength perturbation to a homogeneous

[3]Here we are ignoring real-world selection effects such as magnification bias or evolution bias. Those do vary from one type of tracer to another

FLRW universe. We have presented a gauge invariant construction of the Conformal Fermi Coordinates around a local observer, and have shown that short-scale physical processes experience a modification of the expansion rate H_F and an effective spatial curvature K_F, and by inducing an anisotropic tidal correction to the FLRW metric. Our treatment is linear in the long-wavelength perturbation, but valid on arbitrarily large scales. The range of validity for the CFC is only limited by the scale of the long-wavelength perturbation, but not by the Hubble scale. In the CFC frame, unphysical gauge modes are eliminated by construction.

Using the CFC construction, we have shown that the isotropic effect from a linear long-wavelength perturbation is identical to a modified FLRW cosmology if the long-wavelength perturbation supports no anistropic stress, and that the sound horizon of the problem is either zero, or negligibly small compared to the long wavelength.

In particular, we applied the CFC formalism to study the local influence of a long-wavelength scalar perturbation. We have derived relativistic expressions for the modified expansion rate and the effective spatial curvature as local observables, and found that the tidal potential has an identical form with the Newtonian result on subhorizon scales, as long as the potential in the conformal-Newtonian gauge is used. We then focused on the problem of how short-wavelength, subhorizon density fluctuation evolves in the presence of a long-wavelength perturbation, whose wavelength is not necessarily in the subhorizon regime. Albeit a second-order, general relativistic problem in global coordinates, we have shown that in the CFC frame the long-short coupling can be captured by equations involving linearized Einstein tensor but nonlinear fluid dynamics, resembing the Newtonian theory. The density

field right at the CFC spatial origin can be easily computed, and generalization to a finite vicinity is derived using geodesic deviation. We showed that the locally measured density field modified by the long-short coupling is formally identical to the familiar Standard Perturbation Theory result derived at second order in the Newtonian limit, as long as the long-wavelength density perturbation is evaluated in the synchronous-comoving gauge. We reach the conclusion that for standard single-field inflation, no scale-dependent clustering bias is produced on large scales, even when general relativistic dynamics is included.

Although our analysis is restricted to second order in perturbations, there is no obstacle to generalize to the case of fully nonlinear small-scale modes evolving in the presence of a linear long-wavelength mode.[215] On the other hand, when the long-wavelength mode goes nonlinear, the CFC frame still exists and resembles a uniformly and isotropically expanding spacetime. But departure from the FLRW evolution is implied by the relativistic Raychaudhuri equation, when the long-wavelength perturbation has a non-spherical configuration.[223] Another limitation of the CFC formalism is that it breaks down during the radiation-dominated epoch before recombination, except on scales much larger than the sound horizon of the plasma.

Our analytical work may have promising applications in numerical simulations of structure formation. Through implementation of the CFC frame, the influence of long-wavelength perturbations on short-scale dynamics can be numerically studied using a simulation box much smaller than the wavelength of the long mode, hence saving the huge computational cost of large simulations. It would be interesting to numerically implement the isotropic effect of long-wavelength perturbation for studying the clustering bias of LSS tracers, and

to include the anisotropic tidal correction for studying the tidal alignment of galaxies.

Appendix A

Perturbed Spacetime

Throughout this thesis, we consistently compute up to linear order in metric perturbations. The inverse metric of the perturbed spacetime Eq. (4.1) reads

$$g^{00} = -1 + 2a, \qquad g^{0i} = g^{i0} = b^i, \qquad g^{ij} = \delta^{ij} - h^{ij}. \tag{A.1}$$

Here spatial indices of b_i and h_{ij} can be raised and lowered by δ^{ij} and δ_{ij}. The Christoffel connection coefficients are given by

$$\begin{aligned}
\Gamma^0_{00} &= \dot{a}, && \text{(A.2)}\\
\Gamma^0_{0i} = \Gamma^0_{i0} &= \partial_i a, && \text{(A.3)}\\
\Gamma^0_{ij} &= -\frac{1}{2}\left(\partial_i b_j + \partial_j b_i\right) + \frac{1}{2}\dot{h}_{ij}, && \text{(A.4)}\\
\Gamma^i_{00} &= \partial^i a + \dot{b}^i, && \text{(A.5)}\\
\Gamma^i_{0j} = \Gamma^i_{j0} &= \frac{1}{2}\left(\partial_j b^i - \partial^i b_j\right) + \frac{1}{2}\dot{h}^i{}_j, && \text{(A.6)}\\
\Gamma^k_{ij} &= \frac{1}{2}\left(\partial_i h^k{}_j + \partial_j h^k{}_i - \partial^k h_{ij}\right). && \text{(A.7)}
\end{aligned}$$

Now consider a general linear gauge transformation (a reparametrizaton of the spacetime coordinates)

$$t \longrightarrow t - \xi(\mathbf{x}, t), \qquad x^i \longrightarrow x^i - \xi^i(\mathbf{x}, t). \tag{A.8}$$

The metric perturbations transform as

$$\begin{aligned}
a &\longrightarrow a + \dot{\xi},\\
b_i &\longrightarrow b_i + \dot{\xi}_i - \partial_i \xi,\\
h_{ij} &\longrightarrow h_{ij} + \partial_i \xi_j + \partial_j \xi_i. && \text{(A.9)}
\end{aligned}$$

If we fix the choice for the static observer's tetrads Eqs. (4.2)–(4.3), gauge transforma-

tion Eq. (A.8) corresponds to a new observer having relative velocity $\dot{\xi}^i$ with respect to the original observer. Under this gauge transformation, observer's tetrads undergo a Lorentz transformation, which can be decomposed into a boost β^i and a spatial rotation $\Omega^i{}_j$, i.e.

$$\begin{aligned} e^{\mu}_{(0)} &\rightarrow e^{\mu}_{(0)} + \beta^j\, e^{\mu}_{(j)}, \\ e^{\mu}_{(i)} &\rightarrow \beta_i\, e^{\mu}_{(0)} + \left(\delta_i{}^j + \Omega_i{}^j\right) e^{\mu}_{(j)}. \end{aligned} \tag{A.10}$$

The boost and the rotation are given by

$$\beta^i = \dot{\xi}^i, \qquad \Omega_i{}^j = \frac{1}{2}\left(\partial_i \xi^j - \partial^j \xi_i\right). \tag{A.11}$$

Appendix B

Useful Results for Spherical Harmonics

When computing the lensing kernels, we encounter the following useful integrals involving three spherical harmonic functions:

$$\int d^2n\,[Y_{(lm)}]^* Y^{E,i}_{(J_1M_1)} Y^{E}_{(l_1m_1)i} = -P^{+}_{l_1J_1l}(-)^m\sqrt{\frac{\Pi_{l_1J_1l}}{4\pi}}\begin{pmatrix} l_1 & J_1 & l \\ 1 & -1 & 0\end{pmatrix}\begin{pmatrix} l_1 & J_1 & l \\ m_1 & M_1 & -m\end{pmatrix} \quad \text{(B.1)}$$

$$\int d^2n\,[Y_{(lm)}]^* Y^{B,i}_{(J_1M_1)} Y^{E}_{(l_1m_1)i} = iP^{-}_{l_1J_1l}(-)^m\sqrt{\frac{\Pi_{l_1J_1l}}{4\pi}}\begin{pmatrix} J_1 & l_1 & l \\ 1 & -1 & 0\end{pmatrix}\begin{pmatrix} J_1 & l_1 & l \\ M_1 & m_1 & -m\end{pmatrix} \quad \text{(B.2)}$$

where we have defined the symbol $\Pi_{l_1l_2\cdots} \equiv (2l_1+1)(2l_2+1)\cdots$ and the parity symbol $P^{\pm}_{l_1l_2l_3} \equiv [1\pm(-)^{l_1+l_2+l_3}]/2$. These results can be derived from the results of Ref.[44] Another

integral relevant for computing the polarization is

$$\begin{aligned} & -i\int d^2n\,\left[Y^{s,i''j''}_{(lm)}\right]^*\,Y^{\lambda}_{(JM)i}\,\epsilon^{ijk}\,\hat{n}_j J_k\,Y^{s'}_{(l'm')i''j''} \\ = & -\delta_{ss'}\sqrt{\frac{l'^2-4+l'+\lambda s}{2}}\int d^2n\,\left[{}_{-s}Y_{(lm)}\right]^*\,{}_{-\lambda}Y_{(JM)\,-s'+\lambda}Y_{(l'm')} \\ = & -\delta_{ss'}\sqrt{\frac{l'^2-4+l'+\lambda s}{2}}(-)^m\sqrt{\frac{\Pi_{l'Jl}}{4\pi}}\begin{pmatrix} l & J & l' \\ -s & \lambda & s-\lambda \end{pmatrix}\begin{pmatrix} l & J & l' \\ -m & M & m' \end{pmatrix} \end{aligned} \tag{B.3}$$

where the helicity (s, s', $\lambda, \cdots$) is restricted to ± 2, and J^i is the total angular momentum operator for a spin-two field. The following integrals are relevant for computing polarization rotation

$$\int d^2n\,\left[Y^{TE,ij}_{(lm)}\right]^*\,Y_{(JM)}Y^{TB}_{(l'm')ij} = i(-)^m P^-_{ll'J}\sqrt{\frac{\Pi_{ll'J}}{4\pi}}\begin{pmatrix} l & J & l' \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} l & J & l' \\ -m & M & m' \end{pmatrix} \tag{B.4}$$

$$\int d^2n\,\left[Y^{TB,ij}_{(lm)}\right]^*\,Y_{(JM)}Y^{TE}_{(l'm')ij} = -i(-)^m P^-_{ll'J}\sqrt{\frac{\Pi_{ll'J}}{4\pi}}\begin{pmatrix} l & J & l' \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} l & J & l' \\ -m & M & m' \end{pmatrix} \tag{B.5}$$

$$\int d^2n\,\left[Y^{TE,ij}_{(lm)}\right]^*\,Y_{(JM)}Y^{TE}_{(l'm')ij} = (-)^m P^+_{ll'J}\sqrt{\frac{\Pi_{ll'J}}{4\pi}}\begin{pmatrix} l & J & l' \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} l & J & l' \\ -m & M & m' \end{pmatrix} \tag{B.6}$$

$$\int d^2n\,\left[Y^{TB,ij}_{(lm)}\right]^*\,Y_{(JM)}Y^{TB}_{(l'm')ij} = (-)^m P^+_{ll'J}\sqrt{\frac{\Pi_{ll'J}}{4\pi}}\begin{pmatrix} l & J & l' \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} l & J & l' \\ -m & M & m' \end{pmatrix} \tag{B.7}$$

Completeness relations for spherical harmonic functions are useful. Recall the relation

for the spin-weighted spherical harmonics

$$\sum_M [{}_{s'}Y_{(JM)}(\hat{\mathbf{n}})]^* {}_sY_{(JM)}(\hat{\mathbf{n}}) = \frac{2J+1}{4\pi}\delta_{ss'}, \tag{B.8}$$

for which $s, s' = 0$ would correspond to the ordinary spherical harmonics. Similar relations for vector spherical harmonics read

$$\frac{4\pi}{2J+1}\sum_M [Y^E_{(JM)i}(\hat{\mathbf{n}})]^* Y^E_{(JM)j}(\hat{\mathbf{n}}) = \frac{1}{2}\left(\delta_{ij} - \hat{n}_i\hat{n}_j\right), \tag{B.9}$$

$$\frac{4\pi}{2J+1}\sum_M [Y^B_{(JM)i}(\hat{\mathbf{n}})]^* Y^B_{(JM)j}(\hat{\mathbf{n}}) = \frac{1}{2}\left(\delta_{ij} - \hat{n}_i\hat{n}_j\right), \tag{B.10}$$

$$\frac{4\pi}{2J+1}\sum_M [Y^E_{(JM)i}(\hat{\mathbf{n}})]^* Y^B_{(JM)j}(\hat{\mathbf{n}}) = 0. \tag{B.11}$$

Appendix C

Perturbed FLRW with tensor-scalar coupling

In this Appendix, we collect useful results for a relativistic perturbative approach to solving the tensor-scalar coupling at second order. We consistently derive all expressions to linear order in scalar/matter perturbations, linear order in the tensor perturbation, plus tensor-scalar coupling terms at quadratic order. Note that a few errors in Ref.[59] are corrected here.

C.1 Einstein and fluid equations

First, the metric Eq. (5.1) and its inverse are given by

$$g_{00} = -1 - 2\Phi, \tag{C.1}$$

$$g_{0i} = g_{i0} = a w_i, \tag{C.2}$$

$$g_{ij} = a^2 \left[(1 - 2\Psi)\delta_{ij} + \gamma_{ij} - 2\Psi\gamma_{ij} \right], \tag{C.3}$$

and

$$g^{00} = -1 + 2\Phi, \tag{C.4}$$

$$g^{0i} = g^{i0} = a^{-1} w^i, \tag{C.5}$$

$$g^{ij} = a^{-2} \left[(1 + 2\Psi)\delta^{ij} - \gamma^{ij} - 2\Psi\gamma^{ij} \right]. \tag{C.6}$$

The Christoffel connection coefficients are found to be

$$\begin{aligned}
\Gamma^0_{00} &= \dot{\Phi}, \\
\Gamma^0_{0i} &= \partial_i\Phi + aHw_i, \\
\Gamma^i_{00} &= a^{-2}\left(\delta^{ij} - \gamma^{ij}\right)\partial_j\Phi + a^{-1}\left(\dot{w}^i + Hw^i\right), \\
\Gamma^i_{j0} &= H\delta^i_j + \frac{1}{2}\dot{\gamma}^i_j - \dot{\Psi}\delta^i_j + \frac{1}{2}a^{-1}\left(\partial_j w^i - \partial^i w_j\right), \\
\Gamma^0_{ij} &= a^2\left[H(\delta_{ij}+\gamma_{ij}) + \frac{1}{2}\dot{\gamma}_{ij} - \left(2H(\Psi+\Phi) + \dot{\Psi}\right)(\delta_{ij}+\gamma_{ij}) - (\Psi+\Phi)\dot{\gamma}_{ij}\right] \\
&\quad -\frac{1}{2}a\left(\partial_i w_j + \partial_j w_i\right), \\
\Gamma^k_{ij} &= \frac{1}{2}\left[\partial_i\gamma^k_j + \partial_j\gamma^k_i - \partial^k\gamma_{ij}\right] - \left[\delta^k_j\partial_i\Psi + \delta^k_i\partial_j\Psi - (\delta_{ij}+\gamma_{ij})\left(\delta^{kl}-\gamma^{kl}\right)\partial_l\Psi\right] \\
&\quad +aHw^k\delta_{ij}. \qquad \text{(C.7)}
\end{aligned}$$

so that

$$\Gamma^\mu_{\mu 0} = 3H + \left(-3\dot{\Psi} + \dot{\Phi}\right), \qquad \text{(C.8)}$$

$$\Gamma^\mu_{\mu i} = -3\partial_i\Psi + \partial_i\Phi. \qquad \text{(C.9)}$$

The Einstein tensor $G^{\mu}{}_{\nu} = R^{\mu}{}_{\nu} - (1/2)Rg^{\mu}{}_{\nu}$ has components

$$
\begin{aligned}
G^{0}{}_{0} &= -3H^2 + 6H\dot{\Psi} - 2h^{ij}\partial_i\partial_j\Psi + 6H^2\Phi, \\
G^{0}{}_{i} &= -\left(2\partial_i\dot{\Psi} + 2H\partial_i\Phi\right) + \frac{1}{2}\dot{\gamma}^{j}_{i}\partial_j\left(3\Psi + \Phi\right) + \frac{1}{2a}\partial^2 w_i, \\
G^{i}{}_{0} &= h^{ij}\left(2\partial_j\dot{\Psi} + 2H\partial_j\Phi\right) - \frac{1}{2}a^{-2}\dot{\gamma}^{ij}\partial_j\left(3\Psi + \Phi\right) - \frac{1}{2a^3}\partial^2 w^i - \frac{2}{a}\dot{H}w^i, \\
G^{i}{}_{j} &= -\left(2\dot{H} + 3H^2\right)\delta^{i}_{j} \\
&\quad + \frac{1}{2}\left[\ddot{\gamma}^{i}_{j} + 3H\dot{\gamma}^{i}_{j} - a^{-2}\partial^2\gamma^{i}_{j}\right] \\
&\quad + \left[2\ddot{\Psi} + H\left(6\dot{\Psi} + 2\dot{\Phi}\right) + 2\left(2\dot{H} + 3H^2\right)\Phi - h^{kl}\partial_k\partial_l\left(\Psi - \Phi\right)\right]\delta^{i}_{j} + h^{ik}\partial_k\partial_j\left(\Psi - \Phi\right) \\
&\quad - \Phi\ddot{\gamma}^{i}_{j} - \dot{\gamma}^{i}_{j}\left[3H\Phi + \frac{1}{2}\left(3\dot{\Psi} + \dot{\Phi}\right)\right] - \Psi\partial^2\gamma^{i}_{j} - \frac{1}{2}\left(\partial^i\gamma^{k}_{j} + \partial_j\gamma^{ik} - \partial^k\gamma^{i}_{j}\right)\left(\partial_k\Psi - \partial_k\Phi\right) \\
&\quad - \frac{1}{2a}\left(\partial^i\dot{w}_j + \partial_j\dot{w}^i\right) - \frac{1}{2a}H\left(\partial^i w_j + \partial_j w^i\right),
\end{aligned}
\tag{C.10}
$$

where we define $h_{ij} = a^2(\delta_{ij} + \gamma_{ij})$ and $h^{ij} = a^{-2}(\delta^{ij} - \gamma^{ij})$.

The energy-stress tensor for cold dark matter reads $T^{\mu}{}_{\nu} = \rho_m u^{\mu}u_{\nu}$ where u^{μ} is the 4-velocity. It has components

$$
T^{0}{}_{0} = -\rho_m\left(1 + \delta\right), \quad T^{0}{}_{i} = a\rho_m\left(v_i - w_i\right), \quad T^{i}{}_{0} = -a\rho_m h^{ij}\left(v_j - w_j\right), \quad T^{i}{}_{j} = 0. \tag{C.11}
$$

Using Eq. (C.10) and Eq. (C.11), we can perturbatively expand the Einstein equations $G^{\mu}{}_{\nu} = 8\pi G T^{\mu}{}_{\nu}$, as well as the fluid equations $\nabla_{\mu}T^{\mu}{}_{\nu} = 0$. The time-time, time-space and

space-space Einstein equations are given by

$$h^{ij}\partial_i\partial_j\Psi - 3H\dot{\Psi} - 3H^2\Phi = 4\pi G\rho_m\delta, \tag{C.12}$$

$$-\left(2\partial_i\dot{\Psi} + 2H\partial_i\Phi\right) + \frac{1}{2}\dot{\gamma}_i^j\partial_j\left(3\Psi + \Phi\right) + \frac{1}{2a}\partial^2 w_i = 8\pi G a\rho_m\left(v_i - w_i\right), \tag{C.13}$$

$$\begin{aligned}
&\frac{1}{2}\left[\ddot{\gamma}_j^i + 3H\dot{\gamma}_j^i - a^{-2}\partial^2\gamma_j^i\right] \\
&+\left[2\ddot{\Psi} + H\left(6\dot{\Psi} + 2\dot{\Phi}\right) - h^{kl}\partial_k\partial_l\left(\Psi - \Phi\right)\right]\delta_j^i \\
&+h^{ik}\partial_k\partial_j\left(\Psi - \Phi\right) - \Phi\ddot{\gamma}_j^i - \Psi\partial^2\gamma_j^i \\
&+\dot{\gamma}_j^i\left[H\left(2\Psi - \Phi\right) + \frac{1}{2}\left(\dot{\Psi} - \dot{\Phi}\right)\right] \\
&-\frac{1}{2}\left(\partial^i\gamma_j^k + \partial_j\gamma^{ik} - \partial^k\gamma_j^i\right)\left(\partial_k\Psi - \partial_k\Phi\right) \\
&+\frac{1}{2a}\left(\partial^i\dot{w}_j + \partial_j\dot{w}^i\right) + \frac{1}{2a}H\left(\partial^i w_j + \partial_j w^i\right) = 0.
\end{aligned} \tag{C.14}$$

Note that the homogeneous part has been subtracted. The continuity equation and the Euler equation read

$$\dot{\delta} + ah^{ij}\partial_i v_j - 3\dot{\Psi} = 0, \tag{C.15}$$

$$\dot{v}_i + Hv_i - 2Hw_i + \frac{1}{a}\partial_i\Phi = 0. \tag{C.16}$$

C.2 Solving at second order

To obtain a complete set of equations governing matter perturbations, we collect Eq. (C.12), the divergence of Eq. (C.13), the trace part of Eq. (C.14), Eq. (C.15), and

the curl-free part of Eq. (C.16). Subtracting the first-order part in these equations leads to

$$\frac{\partial^2}{a^2}\Psi^{(2)} - 3H\dot{\Psi}^{(2)} - 3H^2\Phi^{(2)} - 4\pi G\rho_m\delta^{(2)} = \gamma_{ij}\frac{\partial^i\partial^j}{a^2}\Phi^{(1)}, \quad \text{(C.17)}$$

$$-\frac{1}{a}\partial^2\dot{\Psi}^{(2)} - H\frac{1}{a}\partial^2\Phi^{(2)} - 4\pi G\rho_m\partial\cdot v^{(2)} = -\dot{\gamma}^{ij}\frac{1}{a}\partial_i\partial_j\Phi^{(1)}, \quad \text{(C.18)}$$

$$2\ddot{\Psi}^{(2)} + H\left(6\dot{\Psi}^{(2)} + 2\dot{\Phi}^{(2)}\right) + \frac{2}{3}a^{-2}\partial^2\left(-\Psi^{(2)} + \Phi^{(2)}\right) = 0, \quad \text{(C.19)}$$

$$\dot{\delta}^{(2)} + \frac{1}{a}\partial\cdot v^{(2)} - 3\dot{\Psi}^{(2)} = \gamma^{ij}\frac{1}{a}\partial_i v_j^{(1)}, \quad \text{(C.20)}$$

$$\dot{v}_i^{(2)} + Hv_i^{(2)} + \frac{1}{a}\partial_i\Phi^{(2)} = 0. \quad \text{(C.21)}$$

The vector perturbations w_i and $v_{R,i}$ induced at second order, dropping out through this decomposition, are irrelevant for our purpose. At second order, the two scalar potentials are not necessarily identical. We therefore denote $\Gamma^{(2)} = \Psi^{(2)} - \Phi^{(2)}$ to eliminate $\Phi^{(2)}$ in favor of $\Psi^{(2)}$. We find from Eq. (C.19)

$$\ddot{\Psi}^{(2)} + 4H\dot{\Psi}^{(2)} = \frac{1}{3a^2}\partial^2\Gamma^{(2)} + H\dot{\Gamma}^{(2)}. \quad \text{(C.22)}$$

On the other hand, we may rewrite Eq. (C.18) and Eq. (C.21) to find

$$\frac{1}{a}\partial_t\left(a\partial^2\Psi^{(2)}\right) + \frac{3}{2}H^2\left(a\partial\cdot v^{(2)}\right) = \dot{\gamma}^{ij}\partial_i\partial_j\Phi^{(1)} + H\partial^2\Gamma^{(2)}, \quad \text{(C.23)}$$

$$\partial_t\left(a\partial\cdot v^{(2)}\right) + \partial^2\Psi^{(2)} = \partial^2\Gamma^{(2)}. \quad \text{(C.24)}$$

Eliminating the second-order velocity divergence, we derive another equation for the second-order potential

$$\ddot{\Psi}^{(2)} + 4H\dot{\Psi}^{(2)} = \frac{1}{a^2}\partial^{-2}\left[\left(\partial^2\gamma_{ij}\right)\left(\partial^i\partial^j\Phi^{(1)}\right)\right] + H\dot{\Gamma}^{(2)}. \tag{C.25}$$

If Eq. (C.22) is to be consistent with Eq. (C.25), we must have

$$\partial^2\Gamma^{(2)} = 3\partial^{-2}\left[\left(\partial^2\gamma^{ij}\right)\left(\partial_i\partial_j\Phi^{(1)}\right)\right]. \tag{C.26}$$

Once the second-order potential $\Psi^{(2)}$ is found, second-order density and velocity may be found through algebraic relations

$$\delta^{(2)} = \frac{2}{3a^2H^2}\partial^2\Psi^{(2)} - \frac{2}{H}\dot{\Psi}^{(2)} - 2\Psi^{(2)} + 2\Gamma^{(2)} - \frac{2}{3a^2H^2}\gamma^{ij}\partial_i\partial_j\Phi^{(1)}, \tag{C.27}$$

$$\frac{1}{a}\partial\cdot v^{(2)} = -\frac{2}{3a^2H^2}\partial^2\dot{\Psi}^{(2)} - \frac{2}{3a^2H}\partial^2\Psi^{(2)} + \frac{2}{3a^2H}\partial^2\Gamma^{(2)} + \frac{2}{3a^2H^2}\dot{\gamma}^{ij}\partial_i\partial_j\Phi^{(1)}, \tag{C.28}$$

which directly follow from Eq. (C.17) and Eq. (C.18).

Appendix D

Galaxy power quadrupole kernels

The power quadrupole tensor $\mathcal{Q}_{ij}(\mathbf{x})$ is a function of the source location $\mathbf{x}$, which includes a corresponding redshift z and a direction on the sky $\hat{\mathbf{n}}$. The dependence on the sky direction may be expanded using the tensor spherical harmonics of Eq. (2.59),

$$\begin{aligned}\mathcal{Q}_{ij}(\mathbf{x}) &= \gamma^{TE}_{p,JM}(K) \sum_{\alpha=L,VE,TE} \mathcal{Q}^{\alpha}_{J}(K,z) Y^{\alpha}_{(JM)ij}(\hat{\mathbf{n}}) \\ &\quad + \gamma^{TB}_{p,JM}(K,z) \sum_{\alpha=VB,TB} \mathcal{Q}^{\alpha}_{J}(K) Y^{\alpha}_{(JM)ij}(\hat{\mathbf{n}}),\end{aligned} \tag{D.1}$$

The squared sum of the expansion coefficients gives the quantity $\mathcal{Q}^2_J(K,z)$ in Eq. (5.60), i.e.

$$\mathcal{Q}^2_J(K,z) = \sum_{\alpha} \left|\mathcal{Q}^{\alpha}_{J}(K,z)\right|^2, \tag{D.2}$$

for $\alpha = L, VE, VB, TE, TB$. Explicit expressions of $Q_J^\alpha(K,z)$ for a flat FLRW cosmology can be derived using the explicit expressions of the TAM wavefunction (see Ref.[42]). The results are

$$\begin{aligned}
\mathcal{Q}_J^L(K,z) &= -(\kappa_1+\kappa_2)\, j_{J,t}^{(L,TE)}(K\chi) + \\
&\quad \kappa_3\left(\sqrt{\frac{2}{3}}\left(\frac{1}{\chi}-\partial_\chi+a\partial_t\right)\Delta x_\parallel^{TE} + \sqrt{\frac{J(J+1)}{6}}\frac{\Delta x_\perp^{TE}}{\chi}\right), \qquad \text{(D.3)}\\
\mathcal{Q}_J^{VE}(K,z) &= -(\kappa_1+\kappa_2)\, j_{J,t}^{(VE,TE)}(K\chi) \\
&\quad +\kappa_3\left(-\frac{1}{\sqrt{2}}\left(\frac{1}{\chi}-\partial_\chi+a\partial_t\right)\Delta x_\perp^{TE} - \sqrt{\frac{J(J+1)}{2}}\frac{\Delta x_\parallel^{TE}}{\chi}\right), \qquad \text{(D.4)}\\
\mathcal{Q}_J^{TE}(K,z) &= -(\kappa_1+\kappa_2)\, j_{J,t}^{(TE,TE)}(K\chi) + \kappa_3\sqrt{\frac{(J-1)(J+2)}{2}}\frac{\Delta x_\perp^{TE}}{\chi}, \qquad \text{(D.5)}
\end{aligned}$$

from TE-type TAM modes, and

$$\begin{aligned}
\mathcal{Q}_J^{VB}(K) &= -i(\kappa_1+\kappa_2)\, j_{J,t}^{(VB,TB)}(K\chi) + i\kappa_3\left(-\frac{1}{\sqrt{2}}\left[\left(\frac{1}{\chi}-\partial_\chi+a\partial_t\right)\Delta x_\perp^{TB}\right]\right), \qquad \text{(D.6)}\\
\mathcal{Q}_J^{TB}(K) &= -i(\kappa_1+\kappa_2)\, j_{J,t}^{(TB,TB)}(K\chi) + i\kappa_3\sqrt{\frac{(J-1)(J+2)}{2}}\frac{\Delta x_\perp^{TB}}{\chi}, \qquad \text{(D.7)}
\end{aligned}$$

from TB-type TAM modes. Here χ is the comoving radial distance to a source at redshift z. The radial functions $j_{J,t}^{(\alpha,\alpha')}(x)$ can be found in Eqs. (22), (24) and (25) of Ref.[44] Three coefficients have been defined,

$$\begin{aligned}
\kappa_1 &= -\frac{1}{2}\left(\frac{d\ln P_\Phi}{d\ln k} + \mathcal{T}_\gamma(K)\frac{d\ln \mathcal{T}_\delta^2}{d\ln k}\right), \qquad \text{(D.8)}\\
\kappa_2 &= -2\mathcal{S}(K), \qquad \text{(D.9)}\\
\kappa_3 &= -\left(\frac{d\ln P_\Phi}{d\ln k} + \frac{d\ln \mathcal{T}_\delta^2}{d\ln k}\right). \qquad \text{(D.10)}
\end{aligned}$$

Finally, we present expressions for terms involving derivatives of the projection displacement Δx^i. Using Eqs. (5.42)–(5.44), we find

$$\frac{\Delta x_\perp^{TB}}{\chi} = \sqrt{\frac{(J-1)(J+2)}{2}}\mathcal{I}_1, \tag{D.11}$$

$$\left(\frac{1}{\chi} - \partial_\chi + a\partial_t\right)\Delta x_\perp^{TB} = -\sqrt{\frac{(J-1)(J+2)}{2}}\mathcal{T}_\gamma \frac{j_J(K\chi)}{K\chi}, \tag{D.12}$$

$$\frac{\Delta x_\parallel^{TE}}{\chi} = \sqrt{\frac{(J+2)!}{2(J-2)!}}\left(-\frac{1}{2}\frac{1}{K\chi}\mathcal{I}_1 - \frac{1}{2aH\chi}\mathcal{I}_2\right), \tag{D.13}$$

$$\begin{aligned}\left(\frac{1}{\chi} - \partial_\chi + a\partial_t\right)\Delta x_\parallel^{TE} = {} & \sqrt{\frac{(J+2)!}{2(J-2)!}}\frac{1}{2}\left[\left(\mathcal{T}_\gamma + \frac{K}{aH}\frac{\partial\mathcal{T}_\gamma}{\partial(K\tau)}\right)\frac{j_J(K\chi)}{(K\chi)^2}\right. \\ & \left. - \left(\left(1 + \frac{\dot{H}}{H^2}\right) + \frac{1}{aH\chi}\right)\mathcal{I}_2 - \frac{\mathcal{I}_1}{K\chi}\right],\end{aligned} \tag{D.14}$$

$$\frac{\Delta x_\perp^{TE}}{r} = \sqrt{\frac{(J-1)(J+2)}{2}}\left(-\frac{\mathcal{T}_{\gamma,o}}{10}\delta_{J,2} + \mathcal{I}_3\right), \tag{D.15}$$

$$\begin{aligned}\left(\frac{1}{\chi} - \partial_\chi + a\partial_t\right)\Delta x_\perp^{TE} = {} & \sqrt{\frac{(J-1)(J+2)}{2}}\left[-\frac{\mathcal{T}_\gamma}{K\chi}\left(j_J'(K\chi) + \frac{j_J(K\chi)}{K\chi}\right)\right. \\ & \left. + \frac{J(J+1)}{2}\frac{\mathcal{I}_1}{K\chi}\right].\end{aligned} \tag{D.16}$$

Three line-of-sight integrals are needed to describe the projection effects,

$$\mathcal{I}_1 \equiv \int_0^{K\chi} dx\, \mathcal{T}_\gamma(K\tau)\frac{j_J(x)}{x^2}, \tag{D.17}$$

$$\mathcal{I}_2 \equiv \int_0^{K\chi} dx\, \frac{\partial\mathcal{T}_\gamma((K\tau))}{\partial(K\tau)}\frac{j_J(x)}{x^2}, \tag{D.18}$$

$$\mathcal{I}_3 \equiv \int_0^{K\chi} dx\, \frac{\mathcal{T}_\gamma(K\tau)}{x^2}\left(j_J'(x) + \frac{j_J(x)}{x} - \frac{J(J+1)}{2}\left(1 - \frac{x}{Kr}\right)\frac{j_J(x)}{x}\right). \tag{D.19}$$

It is understood that the integrands are evaluated with the past light-cone condition $\chi(z) = \tau_0 - \tau(z)$.

Appendix E

Residual gauge freedom in CFC

For given choice of local expansion history $a_F(\tau_F)$, residual gauge freedom exists in the form of a reparametrization of the spatial coordinate x^i_F at the third order,

$$x^i_F \;\longrightarrow\; x^i_F + \frac{1}{6}A^i{}_{jkl}\,(\tau_F)\; x^j_F x^k_F x^l_F + \mathcal{O}[(x^i_F)^4], \tag{E.1}$$

where the coefficient tensor $A^i{}_{jkl}\,(\tau_F)$ is an arbitrary function of the CFC time and is fully symmetric with respect to its last three indices. This change of coordinates leads to the following change in h^F_{ij}:

$$h^F_{ij} \;\longrightarrow\; h^F_{ij} + A_{(i,j)kl}\, x^k_F x^l_F + \mathcal{O}[(x^i_F)^3]. \tag{E.2}$$

This residual gauge can be used to bring h^F_{ij} into a desirable shape. As an example, one can think of different ways to parametrize spatial hypersurfaces of constant curvature, e.g.

conformally flat and stereographic projections. Note that $A^i{}_{jkl}$ has to drop out of the final result for a physical observable.

Appendix F

The 2nd Friedmann equation in CFC

We verify that the second Friedmann equation holds along the central geodesic in CFC. It follows from Eq. (6.56) that the acceleration is given by,

$$\frac{1}{a_F^2}\frac{d\mathcal{H}_F}{d\tau_F} = \frac{1}{a_F}\frac{d^2 a_F}{dt_F^2} = \frac{1}{a}\frac{d^2 a}{dt^2} - \frac{1}{a^2}\left[\mathcal{H}\Phi' + 2\mathcal{H}'\Phi + \mathcal{H}\Psi' + \Psi'' - \frac{1}{3}\left(\partial\cdot V' + \mathcal{H}\partial\cdot V\right)\right]. \quad \text{(F.1)}$$

We now use the trace (isotropic) part of the ij-Einstein equation, with a general pressure perturbation on the right hand side,

$$2\Psi'' + \mathcal{H}\left(4\Psi' + 2\Phi'\right) + 2\left(2\mathcal{H}' + \mathcal{H}^2\right)\Phi = 8\pi G a^2\left(\mathcal{P}_F - \bar{\mathcal{P}}\right), \quad \text{(F.2)}$$

We then have

$$\begin{aligned} & \mathcal{H}\Phi' + 2\mathcal{H}'\Phi + \mathcal{H}\Psi' + \Psi'' - \frac{1}{3}\left(\partial\cdot V' + \mathcal{H}\partial\cdot V\right) \\ = \;& \frac{1}{3}\left[\partial^2\Psi - 3\mathcal{H}\left(\Psi' + \mathcal{H}\Phi\right)\right] + 4\pi G a^2\left(\mathcal{P}_F - \bar{\mathcal{P}}\right) + \frac{1}{3}\partial^2\left(\Phi - \Psi\right) \\ = \;& \frac{4\pi G}{3}a^2\left[\left(\rho_F - \bar{\rho}\right) + 3\left(\mathcal{P}_F - \bar{\mathcal{P}}\right)\right] + \frac{1}{3}\partial^2\left(\Phi - \Psi\right), \end{aligned} \tag{F.3}$$

by way of the 00-Einstein equation. Again, the second term vanishes if there is no anisotropic stress. Inserting this into Eq. (F.1), we immediately obtain the local version of the second Friedmann equation,

$$\frac{1}{a_F}\frac{d^2 a_F}{dt_F^2} = -\frac{4\pi G}{3}\left(\rho_F + 3\mathcal{P}_F\right), \tag{F.4}$$

where ρ_F and $\mathcal{P}_F$ are the local "background" density and pressure, respectively. They are modified by the underlying long-wavelength perturbation.

Appendix G

Comparison between CFC and SPT

We show that the density corrected for second-order long-short coupling Eq. (6.110) as measured by a local observer is in agreement with the usual result of Standard Perturbation Theory (SPT) at the second order. To compare between the two, we first refer to the same global homogeneous density $\bar{\rho}$ when defining the density constrast δ. This amounts to adding back a contribution $\Delta_{\rm sc}\,\delta^{(1)}$ to Eq. (6.110). Besides, we should adopt the same normalization for the "comoving" scale k by choosing the global scale factor a. From Eq. (6.112), we add back an isotropic displacement term to Eq. (6.110)

$$\left(\frac{a(t_F)}{a_F(t_F)}-1\right)x^i\partial_i\delta^{(1)} \;=\; \frac{1}{3}\Delta_{\rm sc}\,x^i\partial_i\delta^{(1)}. \tag{G.1}$$

After taking into account the above two corrections, we find the second-order correction as a modulation by the long-wavelength density to be given by

$$\delta^{(2)} = \left(\frac{\partial_i\partial_j}{\partial^2}\Delta_{\rm sc}\right) x^j\partial_i\delta^{(1)} + \frac{10}{7}\Delta_{\rm sc}\,\delta^{(1)} + \frac{4}{7}\left(\frac{\partial_i\partial_j}{\partial^2}\Delta_{\rm sc}\right)\left(\frac{\partial^i\partial^j}{\partial^2}\delta^{(1)}\right). \qquad \text{(G.2)}$$

Finally, from the perspective of a "comoving observer" in the fiducial, global cosmology, the CFC defined around the central observer is moving due to the large-scale potential flow. To restore this bulk motion, we may add a homogeneous displacement term (i.e. it is not proportional to x^i), which does not change the power spectrum of the short-scale mode,

$$\left(\frac{\partial_i\partial_j}{\partial^2}\Delta_{\rm sc}\right) x^j\partial_i\delta^{(1)} \;\to\; \left[1 + x^j\partial_j\right]\left(\frac{\partial_i}{\partial^2}\Delta_{\rm sc}\right)\partial_i\delta^{(1)} \to \left(\frac{\partial_i}{\partial^2}\Delta_{\rm sc}\right)\partial_i\delta^{(1)}. \qquad \text{(G.3)}$$

In the first step, we add a bulk displacement term, and in the second step, we adjust the location of the spatial origin $x^i = 0$ to what is chosen in SPT. The final result reads

$$\delta^{(2)} = \left(\frac{\partial_i}{\partial^2}\Delta_{\rm sc}\right)\partial_i\delta^{(1)} + \frac{10}{7}\Delta_{\rm sc}\,\delta^{(1)} + \frac{4}{7}\left(\frac{\partial_i\partial_j}{\partial^2}\Delta_{\rm sc}\right)\left(\frac{\partial^i\partial^j}{\partial^2}\delta^{(1)}\right). \qquad \text{(G.4)}$$

The tensorial structures and numerical coefficients agree completely with the familiar F_2 kernel in second-order SPT computed in EdS universe[224] in the limit that one wave number is much larger than the other, except for an overall factor of two, which accounts for swapping the long-wavelength and short-wavelength modes. While the F_2 kernel is derived from Newtonian nonlinear clustering on subhorizon scales, our CFC result is even valid for (super-)horizon scale k_L. Our result has precisely the same form as the Newtonian result, as long as the long-wavelength density perturbation in the sc gauge is used.

Bibliography

[1] K. S. Dawson *et al.*, "The Baryon Oscillation Spectroscopic Survey of SDSS-III," *Astron.J.*, vol. 145, p. 10, 2013.

[2] S. Alam *et al.*, "The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III," 2015.

[3] U. Seljak *et al.*, "Cosmological parameter analysis including SDSS Ly-alpha forest and galaxy bias: Constraints on the primordial spectrum of fluctuations, neutrino mass, and dark energy," *Phys.Rev.*, vol. D71, p. 103515, 2005.

[4] P. Ade *et al.*, "Planck 2015 results. XV. Gravitational lensing," 2015.

[5] G. Hinshaw *et al.*, "Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results," *Astrophys.J.Suppl.*, vol. 208, p. 19, 2013.

[6] P. Ade *et al.*, "Planck 2015 results. XIII. Cosmological parameters," 2015.

[7] K. Story, C. Reichardt, Z. Hou, R. Keisler, K. Aird *et al.*, "A Measurement of the

Cosmic Microwave Background Damping Tail from the 2500-square-degree SPT-SZ survey," *Astrophys.J.*, vol. 779, p. 86, 2013.

[8] S. Naess *et al.*, "The Atacama Cosmology Telescope: CMB Polarization at $200 < \ell < 9000$," *JCAP*, vol. 1410, no. 10, p. 007, 2014.

[9] A. H. Guth, "Inflationary universe: A possible solution to the horizon and flatness problems," *Phys. Rev. D*, vol. 23, pp. 347–356, Jan 1981. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevD.23.347

[10] K. Sato, "First Order Phase Transition of a Vacuum and Expansion of the Universe," *Mon.Not.Roy.Astron.Soc.*, vol. 195, pp. 467–479, 1981.

[11] A. Albrecht and P. J. Steinhardt, "Cosmology for grand unified theories with radiatively induced symmetry breaking," *Phys. Rev. Lett.*, vol. 48, pp. 1220–1223, Apr 1982. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.48.1220

[12] A. D. Linde, "A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems," *Phys.Lett.*, vol. B108, pp. 389–393, 1982.

[13] A. H. Guth and S.-Y. Pi, "Fluctuations in the new inflationary universe," *Phys. Rev. Lett.*, vol. 49, pp. 1110–1113, Oct 1982. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.49.1110

[14] S. W. Hawking, "The development of irregularities in a single bubble inflationary universe," *Physics Letters B*, vol. 115, pp. 295–297, Sep. 1982.

[15] A. A. Starobinsky, "Dynamics of Phase Transition in the New Inflationary Universe Scenario and Generation of Perturbations," *Phys.Lett.*, vol. B117, pp. 175–178, 1982.

[16] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, "Spontaneous creation of almost scale-free density perturbations in an inflationary universe," *Phys.Rev.*, pp. 679–693, Aug. 1983.

[17] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, "The Effective Field Theory of Inflation," *JHEP*, vol. 0803, p. 014, 2008.

[18] J. Martin, C. Ringeval, and V. Vennin, "Encyclopdia Inflationaris," *Phys.Dark Univ.*, vol. 5-6, p. 75235, 2014.

[19] J. Martin, C. Ringeval, R. Trotta, and V. Vennin, "The Best Inflationary Models After Planck," *JCAP*, vol. 1403, p. 039, 2014.

[20] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," *JHEP*, vol. 0305, p. 013, 2003.

[21] N. Arkani-Hamed and J. Maldacena, "Cosmological Collider Physics," 2015.

[22] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, "The imprints of primordial non-gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects," *Phys.Rev.*, vol. D77, p. 123514, 2008.

[23] F. Schmidt, N. E. Chisari, and C. Dvorkin, "Imprint of inflation on galaxy shape correlations," 2015.

[24] J. Amiaux *et al.*, "Euclid Mission: building of a Reference Survey," *Proc.SPIE Int.Soc.Opt.Eng.*, vol. 8442, p. 0Z, 2012.

[25] J. Green, P. Schechter, C. Baltay, R. Bean, D. Bennett *et al.*, "Wide-Field InfraRed Survey Telescope (WFIRST) Final Report," 2012.

[26] G. Hill, K. Gebhardt, E. Komatsu, N. Drory, P. MacQueen *et al.*, "The Hobby-Eberly Telescope Dark Energy Experiment (HETDEX): Description and Early Pilot Survey Results," *ASP Conf.Ser.*, vol. 399, pp. 115–118, 2008.

[27] R. Maartens, F. B. Abdalla, M. Jarvis, and M. G. Santos, "Overview of Cosmology with the SKA," *PoS*, vol. AASKA14, p. 016, 2015.

[28] J. A. Tyson, "Large synoptic survey telescope: Overview," *Proc.SPIE Int.Soc.Opt.Eng.*, vol. 4836, pp. 10–20, 2002.

[29] O. Dor, J. Bock, P. Capak, R. de Putter, T. Eifler *et al.*, "Cosmology with the SPHEREX All-Sky Spectral Survey," 2014.

[30] L. Abbott and M. B. Wise, "Constraints on Generalized Inflationary Cosmologies," *Nucl.Phys.*, vol. B244, pp. 541–548, 1984.

[31] V. Rubakov, M. Sazhin, and A. Veryaskin, "Graviton Creation in the Inflationary Universe and the Grand Unification Scale," *Phys.Lett.*, vol. B115, pp. 189–192, 1982.

[32] R. Fabbri and M. Pollock, "The Effect of Primordially Produced Gravitons upon the Anisotropy of the Cosmological Microwave Background Radiation," *Phys.Lett.*, vol. B125, pp. 445–448, 1983.

[33] A. A. Starobinsky, "Relict gravitation radiation spectrum and initial state of the universe," *JETP lett*, vol. 30, p. 682, 1979.

[34] P. Ade *et al.*, "Joint Analysis of BICEP2/$KeckArray$ and *Planck* Data," *Phys.Rev.Lett.*, vol. 114, no. 10, p. 101301, 2015.

[35] ——, "Detection of B-Mode Polarization at Degree Angular Scales by BICEP2," *Phys.Rev.Lett.*, vol. 112, no. 24, p. 241101, 2014.

[36] T. Essinger-Hileman, A. Ali, M. Amiri, J. W. Appel, D. Araujo *et al.*, "CLASS: The Cosmology Large Angular Scale Surveyor," *Proc.SPIE Int.Soc.Opt.Eng.*, vol. 9153, p. 91531I, 2014.

[37] Z. Kermish, P. Ade, A. Anthony, K. Arnold, D. Barron *et al.*, "The POLARBEAR Experiment," *Proc.SPIE Int.Soc.Opt.Eng.*, vol. 8452, p. 1C, 2012.

[38] Z. Ahmed *et al.*, "BICEP3: a 95GHz refracting telescope for degree-scale CMB polarization," *Proc.SPIE Int.Soc.Opt.Eng.*, vol. 9153, p. 91531N, 2014.

[39] B. Reichborn-Kjennerud, A. M. Aboobaker, P. Ade, F. Aubin, C. Baccigalupi *et al.*, "EBEX: A balloon-borne CMB polarization experiment," *Proc.SPIE Int.Soc.Opt.Eng.*, vol. 7741, p. 1C, 2010.

[40] A. Kogut, D. Fixsen, D. Chuss, J. Dotson, E. Dwek *et al.*, "The Primordial Inflation Explorer (PIXIE): A Nulling Polarimeter for Cosmic Microwave Background Observations," *JCAP*, vol. 1107, p. 025, 2011.

[41] J. Caligiuri and A. Kosowsky, "Inflationary Tensor Perturbations After BICEP2," *Phys.Rev.Lett.*, vol. 112, p. 191302, 2014.

[42] L. Dai, M. Kamionkowski, and D. Jeong, "Total Angular Momentum Waves for Scalar, Vector, and Tensor Fields," *Phys.Rev.*, vol. D86, p. 125013, 2012.

[43] D. Jeong and M. Kamionkowski, "Clustering Fossils from the Early Universe," *Phys.Rev.Lett.*, vol. 108, p. 251301, 2012.

[44] L. Dai, D. Jeong, and M. Kamionkowski, "Wigner-Eckart theorem in cosmology: Bispectra for total-angular-momentum waves," *Phys.Rev.*, vol. D87, no. 4, p. 043504, 2013.

[45] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, "Quantum Theory of Angular Momentum," 1988.

[46] M. Kamionkowski, A. Kosowsky, and A. Stebbins, "Statistics of cosmic microwave background polarization," *Phys.Rev.*, vol. D55, pp. 7368–7388, 1997.

[47] E. T. Newman and R. Penrose, "Note on the bondi-metzner-sachs group," *Journal of Mathematical Physics*, vol. 7, no. 5, pp. 863–870, 1966. [Online]. Available: http://scitation.aip.org/content/aip/journal/jmp/7/5/10.1063/1.1931221

[48] J. N. Goldberg, A. J. Macfarlane, E. T. Newman, F. Rohrlich, and E. C. G. Sudarshan, "Spins spherical harmonics and ," *Journal of Mathematical Physics*, vol. 8, no. 11, pp. 2155–2161, 1967. [Online]. Available: http://scitation.aip.org/content/aip/journal/jmp/8/11/10.1063/1.1705135

[49] W. Hu and M. J. White, "CMB anisotropies: Total angular momentum method," *Phys.Rev.*, vol. D56, pp. 596–615, 1997.

[50] W. Hu, U. Seljak, M. J. White, and M. Zaldarriaga, "A complete treatment of CMB anisotropies in a FRW universe," *Phys.Rev.*, vol. D57, pp. 3290–3301, 1998.

[51] M. Kamionkowski, L. Dai, and D. Jeong, "Tensor-induced B modes with no temperature fluctuations," *Phys.Rev.*, vol. D89, no. 10, p. 107302, 2014.

[52] K. S. Thorne, "Multipole expansions of gravitational radiation," *Rev. Mod. Phys.*, vol. 52, pp. 299–339, Apr 1980. [Online]. Available: http://link.aps.org/doi/10.1103/RevModPhys.52.299

[53] M. Kamionkowski, A. Kosowsky, and A. Stebbins, "A Probe of primordial gravity waves and vorticity," *Phys.Rev.Lett.*, vol. 78, pp. 2058–2061, 1997.

[54] M. Zaldarriaga and U. Seljak, "An all sky analysis of polarization in the microwave background," *Phys.Rev.*, vol. D55, pp. 1830–1840, 1997.

[55] U. Seljak and M. Zaldarriaga, "Signature of gravity waves in polarization of the microwave background," *Phys.Rev.Lett.*, vol. 78, pp. 2054–2057, 1997.

[56] F. Schmidt and D. Jeong, "Cosmic Rulers," *Phys.Rev.*, vol. D86, p. 083527, 2012.

[57] D. Yamauchi, T. Namikawa, and A. Taruya, "Full-sky formulae for weak lensing power spectra from total angular momentum method," *JCAP*, vol. 1308, p. 051, 2013.

[58] L. Dai, "Rotation of the cosmic microwave background polarization from weak grav-

itational lensing," *Phys.Rev.Lett.*, vol. 112, no. 4, p. 041303, 2014.

[59] L. Dai, D. Jeong, and M. Kamionkowski, "Anisotropic imprint of long-wavelength tensor perturbations on cosmic structure," *Phys.Rev.*, vol. D88, no. 4, p. 043507, 2013.

[60] E. Dimastrogiovanni, M. Fasiello, and M. Kamionkowski, "Imprints of Massive Primordial Fields on Large-Scale Structure," 2015.

[61] E. Dimastrogiovanni, M. Fasiello, D. Jeong, and M. Kamionkowski, "Inflationary tensor fossils in large-scale structure," *JCAP*, vol. 1412, no. 12, p. 050, 2014.

[62] L. Dai, D. Jeong, and M. Kamionkowski, "Seeking Inflation Fossils in the Cosmic Microwave Background," *Phys.Rev.*, vol. D87, no. 10, p. 103006, 2013.

[63] S. K. Lee, "Harmonics in the Dark-Matter Sky: Directional Detection in the Fourier-Bessel Basis," *JCAP*, vol. 1403, p. 047, 2014.

[64] P. Ade *et al.*, "Planck 2013 results. XVI. Cosmological parameters," *Astron.Astrophys.*, vol. 571, p. A16, 2014.

[65] J. Martin, C. Ringeval, and V. Vennin, "Encyclopdia Inflationaris," *Phys.Dark Univ.*, 2014.

[66] A. D. Linde, "Hybrid inflation," *Phys.Rev.*, vol. D49, pp. 748–754, 1994.

[67] E. J. Copeland, A. R. Liddle, D. H. Lyth, E. D. Stewart, and D. Wands, "False vacuum inflation with Einstein gravity," *Phys.Rev.*, vol. D49, pp. 6410–6433, 1994.

[68] A. R. Liddle, A. Mazumdar, and F. E. Schunck, "Assisted inflation," *Phys.Rev.*, vol. D58, p. 061301, 1998.

[69] D. H. Lyth and D. Wands, "Generating the curvature perturbation without an inflaton," *Phys.Lett.*, vol. B524, pp. 5–14, 2002.

[70] D. Baumann and D. Green, "Signatures of Supersymmetry from the Early Universe," *Phys.Rev.*, vol. D85, p. 103520, 2012.

[71] X. Chen and Y. Wang, "Quasi-Single Field Inflation and Non-Gaussianities," *JCAP*, vol. 1004, p. 027, 2010.

[72] K. Dimopoulos, M. Karciauskas, D. H. Lyth, and Y. Rodriguez, "Statistical anisotropy of the curvature perturbation from vector field perturbations," *JCAP*, vol. 0905, p. 013, 2009.

[73] A. Golovnev and V. Vanchurin, "Cosmological perturbations from vector inflation," *Phys.Rev.*, vol. D79, p. 103524, 2009.

[74] N. Bartolo, E. Dimastrogiovanni, S. Matarrese, and A. Riotto, "Anisotropic bispectrum of curvature perturbations from primordial non-Abelian vector fields," *JCAP*, vol. 0910, p. 015, 2009.

[75] J. Beltran Jimenez and A. L. Maroto, "Cosmological evolution in vector-tensor theories of gravity," *Phys.Rev.*, vol. D80, p. 063512, 2009.

[76] D. Seery, M. S. Sloth, and F. Vernizzi, "Inflationary trispectrum from graviton exchange," *JCAP*, vol. 0903, p. 018, 2009.

[77] S. B. Giddings and M. S. Sloth, "Semiclassical relations and IR effects in de Sitter and slow-roll space-times," *JCAP*, vol. 1101, p. 023, 2011.

[78] ——, "Cosmological observables, IR growth of fluctuations, and scale-dependent anisotropies," *Phys.Rev.*, vol. D84, p. 063528, 2011.

[79] X. Chen, "Primordial Non-Gaussianities from Inflation Models," *Adv.Astron.*, vol. 2010, p. 638979, 2010.

[80] W. P. Elderton, "Tables for testing the goodness of fit of theory to observation." *Biometrika*, vol. 1, no. 2, pp. 155–163, 1902. [Online]. Available: http://biomet.oxfordjournals.org/content/1/2/155.short

[81] M. Akhshik, "Clustering Fossils in Solid Inflation," 2014.

[82] P. Ade *et al.*, "Planck 2013 results. XV. CMB power spectra and likelihood," *Astron.Astrophys.*, vol. 571, p. A15, 2014.

[83] ——, "Planck 2015 results. XVII. Constraints on primordial non-Gaussianity," 2015.

[84] A. Hajian, T. Souradeep, and N. J. Cornish, "Statistical isotropy of the WMAP data: A Bipolar power spectrum analysis," *Astrophys.J.*, vol. 618, pp. L63–L66, 2004.

[85] M. Kamionkowski and T. Souradeep, "The Odd-Parity CMB Bispectrum," *Phys.Rev.*, vol. D83, p. 027301, 2011.

[86] L. G. Book, M. Kamionkowski, and T. Souradeep, "Odd-Parity Bipolar Spherical Harmonics," *Phys.Rev.*, vol. D85, p. 023010, 2012.

[87] L. Book, M. Kamionkowski, and F. Schmidt, "Lensing of 21-cm Fluctuations by Primordial Gravitational Waves," *Phys.Rev.Lett.*, vol. 108, p. 211301, 2012.

[88] E. Komatsu and D. N. Spergel, "Acoustic signatures in the primary microwave background bispectrum," *Phys.Rev.*, vol. D63, p. 063002, 2001.

[89] A. Lewis and A. Challinor, "Weak gravitational lensing of the cmb," *Phys.Rept.*, vol. 429, pp. 1–65, 2006.

[90] W. Hu, "Weak lensing of the CMB: A harmonic approach," *Phys.Rev.*, vol. D62, p. 043007, 2000.

[91] K. M. Smith, O. Zahn, and O. Dore, "Detection of Gravitational Lensing in the Cosmic Microwave Background," *Phys.Rev.*, vol. D76, p. 043510, 2007.

[92] C. M. Hirata, S. Ho, N. Padmanabhan, U. Seljak, and N. A. Bahcall, "Correlation of CMB with large-scale structure: II. Weak lensing," *Phys.Rev.*, vol. D78, p. 043520, 2008.

[93] D. Hanson *et al.*, "Detection of B-mode Polarization in the Cosmic Microwave Background with Data from the South Pole Telescope," *Phys.Rev.Lett.*, vol. 111, no. 14, p. 141301, 2013.

[94] P. Ade *et al.*, "A Measurement of the Cosmic Microwave Background B-Mode Polarization Power Spectrum at Sub-Degree Scales with POLARBEAR," *Astrophys.J.*, vol. 794, no. 2, p. 171, 2014.

[95] M. Kamionkowski, T. L. Smith, and A. Heavens, "The CMB Bispectrum, Trispec-

trum, non-Gaussianity, and the Cramer-Rao Bound," *Phys.Rev.*, vol. D83, p. 023007, 2011.

[96] E. Komatsu *et al.*, "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," *Astrophys.J.Suppl.*, vol. 192, p. 18, 2011.

[97] A. Lewis, A. Challinor, and A. Lasenby, "Efficient computation of CMB anisotropies in closed FRW models," *Astrophys.J.*, vol. 538, pp. 473–476, 2000.

[98] E. Nelson and S. Shandera, "Statistical Naturalness and non-Gaussianity in a Finite Universe," *Phys.Rev.Lett.*, vol. 110, no. 13, p. 131301, 2013.

[99] B. Bayta, A. Kesavan, E. Nelson, S. Park, and S. Shandera, "Nonlocal bispectra from super cosmic variance," *Phys.Rev.*, vol. D91, no. 8, p. 083518, 2015.

[100] N. Kogo and E. Komatsu, "Angular trispectrum of cmb temperature anisotropy from primordial non-gaussianity with the full radiation transfer function," *Phys.Rev.*, vol. D73, p. 083007, 2006.

[101] J. Smidt, A. Amblard, C. T. Byrnes, A. Cooray, A. Heavens *et al.*, "CMB Constraints on Primordial non-Gaussianity from the Bispectrum (f_{NL}) and Trispectrum (g_{NL} and τ_{NL}) and a New Consistency Test of Single-Field Inflation," *Phys.Rev.*, vol. D81, p. 123007, 2010.

[102] L. Senatore and M. Zaldarriaga, "The Effective Field Theory of Multifield Inflation," *JHEP*, vol. 1204, p. 024, 2012.

[103] X. Chen, B. Hu, M.-x. Huang, G. Shiu, and Y. Wang, "Large Primordial Trispectra in General Single Field Inflation," *JCAP*, vol. 0908, p. 008, 2009.

[104] F. Arroja, S. Mizuno, K. Koyama, and T. Tanaka, "On the full trispectrum in single field DBI-inflation," *Phys.Rev.*, vol. D80, p. 043527, 2009.

[105] N. Bartolo, M. Fasiello, S. Matarrese, and A. Riotto, "Large non-Gaussianities in the Effective Field Theory Approach to Single-Field Inflation: the Trispectrum," *JCAP*, vol. 1009, p. 035, 2010.

[106] S. Basak, A. Hajian, and T. Souradeep, "Statistical isotropy of cmb polarization maps," *Phys.Rev.*, vol. D74, p. 021301, 2006.

[107] A. R. Pullen and M. Kamionkowski, "Cosmic Microwave Background Statistics for a Direction-Dependent Primordial Power Spectrum," *Phys.Rev.*, vol. D76, p. 103529, 2007.

[108] A. Challinor and F. van Leeuwen, "Peculiar velocity effects in high resolution microwave background experiments," *Phys.Rev.*, vol. D65, p. 103001, 2002.

[109] R. K. Sachs and A. M. Wolfe, "Perturbations of a Cosmological Model and Angular Variations of the Microwave Background," *APJ*, vol. 147, p. 73, Jan. 1967.

[110] M. J. Rees and D. W. Sciama, "Large-scale Density Inhomogeneities in the Universe," *Nature*, vol. 217, pp. 511–516, Feb. 1968.

[111] P. Ade *et al.*, "Planck 2015 results. XXI. The integrated Sachs-Wolfe effect," 2015.

[112] A. Blanchard and J. Schneider, "Gravitational lensing effect on the fluctuations of the cosmic background radiation," *AAP*, vol. 184, pp. 1–6, Oct. 1987.

[113] A. Kashlinsky, "Small-scale fluctuations in the microwave background radiation and multiple gravitational lensing," *APJ Letter*, vol. 331, pp. L1–L4, Aug. 1988.

[114] U. Seljak, "Gravitational lensing effect on cosmic microwave background anisotropies: A Power spectrum approach," *Astrophys.J.*, vol. 463, p. 1, 1996.

[115] M. Zaldarriaga and U. c. v. Seljak, "Reconstructing projected matter density power spectrum from cosmic microwave background," *Phys. Rev. D*, vol. 59, p. 123507, May 1999. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevD.59.123507

[116] D. Yamauchi, T. Namikawa, and A. Taruya, "Weak lensing generated by vector perturbations and detectability of cosmic strings," *JCAP*, vol. 1210, p. 030, 2012.

[117] S. Mollerach, "Gravitational lensing on the cosmic microwave background by gravity waves," *Phys.Rev.*, vol. D57, pp. 1303–1305, 1998.

[118] A. Cooray, M. Kamionkowski, and R. R. Caldwell, "Cosmic shear of the microwave background: The Curl diagnostic," *Phys.Rev.*, vol. D71, p. 123527, 2005.

[119] C. Li and A. Cooray, "Weak Lensing of the Cosmic Microwave Background by Foreground Gravitational Waves," *Phys.Rev.*, vol. D74, p. 023521, 2006.

[120] M. Zaldarriaga and U. Seljak, "Gravitational lensing effect on cosmic microwave background polarization," *Phys.Rev.*, vol. D58, p. 023003, 1998.

[121] L. G. Book and E. E. Flanagan, "Astrometric Effects of a Stochastic Gravitational Wave Background," *Phys.Rev.*, vol. D83, p. 024024, 2011.

[122] C. Pitrou, "The Radiative transfer at second order: A Full treatment of the Boltzmann equation with polarization," *Class.Quant.Grav.*, vol. 26, p. 065006, 2009.

[123] B. Mashhoon, "Influence of gravitation on the propagation of electromagnetic radiation," *Phys. Rev. D*, vol. 11, pp. 2679–2684, May 1975. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevD.11.2679

[124] H. Ishihara, M. Takahashi, and A. Tomimatsu, "Gravitational faraday rotation induced by a kerr black hole," *Phys. Rev. D*, vol. 38, pp. 472–477, Jul 1988. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevD.38.472

[125] M. Nouri-Zonoz, "Gravitoelectromagnetic approach to the gravitational faraday rotation in stationary spacetimes," *Phys. Rev. D*, vol. 60, p. 024013, Jun 1999. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevD.60.024013

[126] S. Kopeikin and B. Mashhoon, "Gravitomagnetic effects in the propagation of electromagnetic waves in variable gravitational fields of arbitrary moving and spinning bodies," *Phys.Rev.*, vol. D65, p. 064025, 2002.

[127] M. Sereno, "Gravitational Faraday rotation in a weak gravitational field," *Phys.Rev.*, vol. D69, p. 087501, 2004.

[128] P. Nag, S. Bharadwaj, and S. Kar, "Can the rotation of the dark matter halo of our

Galaxy be detected through its effect on the cosmic microwave background polarisation?" *JCAP*, vol. 0507, p. 006, 2005.

[129] D. Baumann, "TASI Lectures on Inflation," 2009.

[130] A. Challinor and G. Chon, "Geometry of weak lensing of CMB polarization," *Phys.Rev.*, vol. D66, p. 127301, 2002.

[131] A. Challinor and A. Lewis, "Lensed CMB power spectra from all-sky correlation functions," *Phys.Rev.*, vol. D71, p. 103010, 2005.

[132] N. Kaiser and J. A. Peacock, "On the Bias of the Distance-Redshift Relation from Gravitational Lensing," 2015.

[133] A. Rotti and T. Souradeep, "A New Window into Stochastic Gravitational Wave Background," *Phys.Rev.Lett.*, vol. 109, p. 221301, 2012.

[134] D. Sarkar, P. Serra, A. Cooray, K. Ichiki, and D. Baumann, "Cosmic shear from scalar-induced gravitational waves," *Phys.Rev.*, vol. D77, p. 103515, 2008.

[135] H. Padmanabhan, A. Rotti, and T. Souradeep, "A comparison of CMB lensing efficiency of gravitational waves and large scale structure," *Phys.Rev.*, vol. D88, p. 063507, 2013.

[136] T. Souradeep, *private communication*.

[137] T. Namikawa, "Cosmology from weak lensing of CMB," *PTEP*, vol. 2014, no. 6, p. 06B108, 2014.

[138] T. Namikawa, D. Yamauchi, and A. Taruya, "Full-sky lensing reconstruction of gradient and curl modes from CMB maps," *JCAP*, vol. 1201, p. 007, 2012.

[139] ——, "Constraining cosmic string parameters with curl mode of CMB lensing," *Phys.Rev.*, vol. D88, no. 8, p. 083525, 2013.

[140] M. Li and X. Zhang, "Cosmological CPT violating effect on CMB polarization," *Phys.Rev.*, vol. D78, p. 103516, 2008.

[141] M. Kamionkowski, "How to De-Rotate the Cosmic Microwave Background Polarization," *Phys.Rev.Lett.*, vol. 102, p. 111302, 2009.

[142] V. Gluscevic, M. Kamionkowski, and A. Cooray, "De-Rotation of the Cosmic Microwave Background Polarization: Full-Sky Formalism," *Phys.Rev.*, vol. D80, p. 023510, 2009.

[143] A. P. Yadav, . Biswas, Rahul, M. Su, and M. Zaldarriaga, "Constraining a spatially dependent rotation of the Cosmic Microwave Background Polarization," *Phys.Rev.*, vol. D79, p. 123009, 2009.

[144] A. Cooray and W. Hu, "Second order corrections to weak lensing by large scale structure," *Astrophys.J.*, vol. 574, p. 19, 2002.

[145] C. M. Hirata and U. Seljak, "Reconstruction of lensing from the cosmic microwave background polarization," *Phys.Rev.*, vol. D68, p. 083002, 2003.

[146] U.-L. Pen and S. Mao, "Rotation in gravitational lenses," *Mon.Not.Roy.Astron.Soc.*, vol. 367, pp. 1543–1550, 2006.

[147] C. Shapiro and A. Cooray, "The born and lens-lens corrections to weak gravitational lensing angular power spectra," *JCAP*, vol. 0603, p. 007, 2006.

[148] E. Krause and C. M. Hirata, "Weak lensing power spectra for precision cosmology: Multiple-deflection, reduced shear and lensing bias corrections," *Astron.Astrophys.*, vol. 523, p. A28, 2010.

[149] U. Seljak, *private communication*.

[150] T. L. Smith, M. Kamionkowski, and A. Cooray, "Direct detection of the inflationary gravitational wave background," *Phys.Rev.*, vol. D73, p. 023504, 2006.

[151] S. Chongchitnan and G. Efstathiou, "Prospects for direct detection of primordial gravitational waves," *Phys.Rev.*, vol. D73, p. 083511, 2006.

[152] T. L. Smith, M. Kamionkowski, and A. Cooray, "The inflationary gravitational-wave background and measurements of the scalar spectral index," *Phys.Rev.*, vol. D78, p. 083525, 2008.

[153] J. Caligiuri, A. Kosowsky, W. H. Kinney, and N. Seto, "The History of Inflation from Microwave Background Polarimetry and Laser Interferometry," 2014.

[154] D. Jeong and F. Schmidt, "Large-Scale Structure with Gravitational Waves I: Galaxy Clustering," *Phys.Rev.*, vol. D86, p. 083512, 2012.

[155] F. Schmidt and D. Jeong, "Large-Scale Structure with Gravitational Waves II: Shear," *Phys.Rev.*, vol. D86, p. 083513, 2012.

[156] S. Dodelson, E. Rozo, and A. Stebbins, "Primordial gravity waves and weak lensing," *Phys.Rev.Lett.*, vol. 91, p. 021301, 2003.

[157] U.-L. Pen, "Gravitational lensing of pre-reionization gas," *New Astron.*, vol. 9, pp. 417–424, 2004.

[158] N. Kaiser and A. H. Jaffe, "Bending of light by gravity waves," *Astrophys.J.*, vol. 484, pp. 545–554, 1997.

[159] K. W. Masui and U.-L. Pen, "Primordial gravity wave fossils and their use in testing inflation," *Phys.Rev.Lett.*, vol. 105, p. 161302, 2010.

[160] E. Bertschinger, "Cosmological dynamics: Course 1," 1993.

[161] P. Creminelli and M. Zaldarriaga, "Single field consistency relation for the 3-point function," *JCAP*, vol. 0410, p. 006, 2004.

[162] L. Dai, E. Pajer, and F. Schmidt, "On Separate Universes," 2015.

[163] R. de Putter, O. Dor, and D. Green, "Is There Scale-Dependent Bias in Single-Field Inflation?" 2015.

[164] K. S. Thorne, "Gravitational-wave bursts with memory: The christodoulou effect," *Phys. Rev. D*, vol. 45, pp. 520–524, Jan 1992. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevD.45.520

[165] D. Christodoulou, "Nonlinear nature of gravitation and gravitational wave experiments," *Phys.Rev.Lett.*, vol. 67, pp. 1486–1489, 1991.

[166] R. van Haasteren and Y. Levin, "Gravitational-wave memory and pulsar timing arrays," *Mon.Not.Roy.Astron.Soc.*, vol. 401, p. 2372, 2010.

[167] M. Favata, "The gravitational-wave memory effect," *Class.Quant.Grav.*, vol. 27, p. 084036, 2010.

[168] D. Jeong, F. Schmidt, and C. M. Hirata, "Large-scale clustering of galaxies in general relativity," *Phys.Rev.*, vol. D85, p. 023504, 2012.

[169] J. Yoo, N. Hamaus, U. Seljak, and M. Zaldarriaga, "Going beyond the Kaiser redshift-space distortion formula: a full general relativistic account of the effects and their detectability in galaxy clustering," *Phys.Rev.*, vol. D86, p. 063514, 2012.

[170] D. Jeong and F. Schmidt, "Cosmic Clocks," *Phys.Rev.*, vol. D89, no. 4, p. 043519, 2014.

[171] L. Lombriser, J. Yoo, and K. Koyama, "Relativistic effects in galaxy clustering in a parametrized post-Friedmann universe," *Phys.Rev.*, vol. D87, p. 104019, 2013.

[172] D. Bertacca, R. Maartens, and C. Clarkson, "Observed galaxy number counts on the lightcone up to second order: I. Main result," *JCAP*, vol. 1409, no. 09, p. 037, 2014.

[173] D. Jeong and F. Schmidt, "Large-Scale Structure Observables in General Relativity," *Class.Quant.Grav.*, vol. 32, no. 4, p. 044001, 2015.

[174] E. Pajer, F. Schmidt, and M. Zaldarriaga, "The Observed Squeezed Limit of Cosmological Three-Point Functions," *Phys.Rev.*, vol. D88, no. 8, p. 083502, 2013.

[175] J. Yoo, "Proper-time hypersurface of nonrelativistic matter flows: Galaxy bias in general relativity," *Phys.Rev.*, vol. D90, no. 12, p. 123507, 2014.

[176] J. Yoo and M. Zaldarriaga, "Beyond the Linear-Order Relativistic Effect in Galaxy Clustering: Second-Order Gauge-Invariant Formalism," *Phys.Rev.*, vol. D90, no. 2, p. 023513, 2014.

[177] S. Brahma, E. Nelson, and S. Shandera, "Fossilized Gravitational Wave Relic and Primordial Clocks," *Phys.Rev.*, vol. D89, no. 2, p. 023507, 2014.

[178] F. Schmidt, E. Pajer, and M. Zaldarriaga, "Large-Scale Structure and Gravitational Waves III: Tidal Effects," *Phys.Rev.*, vol. D89, no. 8, p. 083507, 2014.

[179] N. Kaiser, "Clustering in real space and in redshift space," *Mon.Not.Roy.Astron.Soc.*, vol. 227, pp. 1–27, 1987.

[180] R. Mandelbaum, C. M. Hirata, M. Ishak, U. Seljak, and J. Brinkmann, "Detection of large scale intrinsic ellipticity-density correlation from the sloan digital sky survey and implications for weak lensing surveys," *Mon.Not.Roy.Astron.Soc.*, vol. 367, pp. 611–626, 2006.

[181] C. M. Hirata, R. Mandelbaum, M. Ishak, U. Seljak, R. Nichol *et al.*, "Intrinsic galaxy alignments from the 2SLAQ and SDSS surveys: Luminosity and redshift scalings and implications for weak lensing surveys," *Mon.Not.Roy.Astron.Soc.*, vol. 381, pp. 1197–1218, 2007.

[182] T. Okumura, Y. Jing, and C. Li, "Intrinsic Ellipticity Correlation of SDSS Luminous

Red Galaxies and Misalignment with their Host Dark Matter Halos," *Astrophys.J.*, vol. 694, pp. 214–221, 2009.

[183] T. Okumura and Y. Jing, "The Gravitational Shear – Intrinsic Ellipticity Correlation Functions of Luminous Red Galaxies in Observation and in ΛCDM model," *Astrophys.J.*, vol. 694, pp. L83–L86, 2009.

[184] B. Joachimi, R. Mandelbaum, F. Abdalla, and S. Bridle, "Constraints on intrinsic alignment contamination of weak lensing surveys using the MegaZ-LRG sample," *Astron.Astrophys.*, vol. 527, p. A26, 2011.

[185] R. A. Croft and C. A. Metzler, "Weak lensing surveys and the intrinsic correlation of galaxy ellipticities," *Astrophys.J.*, vol. 545, pp. 561–571, 2000.

[186] A. Heavens, A. Refregier, and C. Heymans, "Intrinsic correlation of galaxy shapes: Implications for weak lensing measurements," *Mon.Not.Roy.Astron.Soc.*, vol. 319, p. 649, 2000.

[187] J. Lee and U.-L. Pen, "Cosmic shear from galaxy spins," *Astrophys.J.*, vol. 532, p. L5, 2000.

[188] ——, "Galaxy spin statistics and spin-density correlation," *Astrophys.J.*, vol. 555, pp. 106–124, 2001.

[189] P. Catelan, M. Kamionkowski, and R. D. Blandford, "Intrinsic and extrinsic galaxy alignment," *Mon.Not.Roy.Astron.Soc.*, vol. 320, pp. L7–L13, 2001.

[190] L. Hui and J. Zhang, "Intrinsic/extrinsic density-ellipticity correlations and galaxy-galaxy lensing," 2002.

[191] C. M. Hirata and U. Seljak, "Intrinsic alignment-lensing interference as a contaminant of cosmic shear," *Phys.Rev.*, vol. D70, p. 063526, 2004.

[192] M. D. Schneider and S. Bridle, "A halo model for intrinsic alignments of galaxy ellipticities," *Mon.Not.Roy.Astron.Soc.*, vol. 402, p. 2127, 2010.

[193] M. D. Schneider, C. S. Frenk, and S. Cole, "The Shapes and Alignments of Dark Matter Halos," *JCAP*, vol. 1205, p. 030, 2012.

[194] J. Blazek, M. McQuinn, and U. Seljak, "Testing the tidal alignment model of galaxy intrinsic alignment," *JCAP*, vol. 1105, p. 010, 2011.

[195] R. Laureijs *et al.*, "Euclid Definition Study Report," 2011.

[196] Z. Ivezic, J. Tyson, R. Allsman, J. Andrew, and R. Angel, "LSST: from Science Drivers to Reference Design and Anticipated Data Products," 2008.

[197] N. E. Chisari, C. Dvorkin, and F. Schmidt, "Can weak lensing surveys confirm BICEP2?" *Phys.Rev.*, vol. D90, no. 4, p. 043527, 2014.

[198] T. Baldauf, U. Seljak, L. Senatore, and M. Zaldarriaga, "Galaxy Bias and non-Linear Structure Formation in General Relativity," *JCAP*, vol. 1110, p. 031, 2011.

[199] F. Schmidt, D. Jeong, and V. Desjacques, "Peak-Background Split, Renormalization, and Galaxy Clustering," *Phys.Rev.*, vol. D88, no. 2, p. 023515, 2013.

[200] P. McDonald, "Toward a measurement of the cosmological geometry at Z 2: predicting lyman-alpha forest correlation in three dimensions, and the potential of future data sets," *Astrophys.J.*, vol. 585, pp. 34–51, 2003.

[201] E. Sirko, "Initial conditions to cosmological N-body simulations, or how to run an ensemble of simulations," *Astrophys.J.*, vol. 634, pp. 728–743, 2005.

[202] B. D. Sherwin and M. Zaldarriaga, "The Shift of the Baryon Acoustic Oscillation Scale: A Simple Physical Picture," *Phys.Rev.*, vol. D85, p. 103523, 2012.

[203] Y. Li, W. Hu, and M. Takada, "Super-Sample Covariance in Simulations," *Phys.Rev.*, vol. D89, no. 8, p. 083519, 2014.

[204] C. Wagner, F. Schmidt, C.-T. Chiang, and E. Komatsu, "Separate Universe Simulations," *Mon.Not.Roy.Astron.Soc.*, vol. 448, p. 11, 2015.

[205] S. Deser and J. Franklin, "Schwarzschild and Birkhoff a la Weyl," *Am.J.Phys.*, vol. 73, pp. 261–264, 2005.

[206] J. E. Gunn, "On the Propagation of Light in Inhomogeneous Cosmologies. I. Mean Effects," *APJ*, vol. 150, p. 737, Dec. 1967.

[207] R. Kantowski, "Corrections in the Luminosity-Redshift Relations of the Homogeneous Fried-Mann Models," *APJ*, vol. 155, p. 89, Jan. 1969.

[208] C. C. Dyer and R. C. Roeder, "Observations in Locally Inhomogeneous Cosmological Models," *APJ*, vol. 189, pp. 167–176, Apr. 1974.

[209] F. K. Manasse and C. W. Misner, "Fermi normal coordinates and some basic concepts in differential geometry," *Journal of Mathematical Physics*, vol. 4, no. 6, pp. 735–745, 1963. [Online]. Available: http://scitation.aip.org/content/aip/journal/jmp/4/6/10.1063/1.1724316

[210] L. Dai, E. Pajer, and F. Schmidt, "Conformal Fermi Coordinates," 2015.

[211] S. Carroll, "Spacetime and geometry: An introduction to general relativity," 2003.

[212] R. Bean and O. Dore, "Probing dark energy perturbations: The Dark energy equation of state and speed of sound as measured by WMAP," *Phys.Rev.*, vol. D69, p. 083503, 2004.

[213] P. Creminelli, G. D'Amico, J. Norena, L. Senatore, and F. Vernizzi, "Spherical collapse in quintessence models with zero speed of sound," *JCAP*, vol. 1003, p. 027, 2010.

[214] D. H. Lyth, K. A. Malik, and M. Sasaki, "A General proof of the conservation of the curvature perturbation," *JCAP*, vol. 0505, p. 004, 2005.

[215] C. Wagner, F. Schmidt, C.-T. Chiang, and E. Komatsu, "The angle-averaged squeezed limit of nonlinear matter N-point functions," 2015.

[216] N. Kaiser, "On the Spatial correlations of Abell clusters," *Astrophys.J.*, vol. 284, pp. L9–L12, 1984.

[217] M. Alvarez, T. Baldauf, J. R. Bond, N. Dalal, R. de Putter *et al.*, "Testing Inflation with Large Scale Structure: Connecting Hopes with Reality," 2014.

[218] N. Bartolo, S. Matarrese, and A. Riotto, "Signatures of primordial non-Gaussianity in the large-scale structure of the Universe," *JCAP*, vol. 0510, p. 010, 2005.

[219] L. Verde and S. Matarrese, "Detectability of the effect of Inflationary non-Gaussianity on halo bias," *Astrophys.J.*, vol. 706, pp. L91–L95, 2009.

[220] M. Bruni, J. C. Hidalgo, N. Meures, and D. Wands, "Non-Gaussian Initial Conditions in CDM: Newtonian, Relativistic, and Primordial Contributions," *Astrophys.J.*, vol. 785, p. 2, 2014.

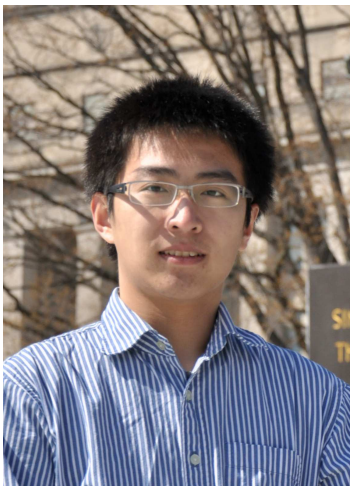
[221] D. Bertacca, N. Bartolo, M. Bruni, K. Koyama, R. Maartens *et al.*, "Galaxy bias and gauges at second order in General Relativity," 2015.

[222] N. Bartolo, D. Bertacca, M. Bruni, K. Koyama, R. Maartens *et al.*, "A relativistic signature in large-scale structure: Scale-dependent bias from single-field inflation," 2015.

[223] A. Raychaudhuri, "Relativistic cosmology. i," *Phys. Rev.*, vol. 98, pp. 1123–1126, May 1955. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRev.98.1123

[224] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, "Large scale structure of the universe and cosmological perturbation theory," *Phys.Rept.*, vol. 367, pp. 1–248, 2002.

Vita



Liang Dai was born on August 4th, 1988 in Hangzhou in the eastern China. Since 2007 he became a physics-major undergraduate in Peking University, Beijing, and worked on perturbative quantum chromodynamics under the supervision of Prof. Chongsheng Li. After receiving his Bachelor's degree in 2011, Dai became a graduate student with Prof. Marc Kamionkowski at Johns Hopkins University and conducted research in cosmology. He received a Ph.D. degree in physics in 2015 for his dissertation on cosmological perturbation. Dai was awarded the Dr. Pliny A. and Margaret H. Price Prize in 2013. Starting in 2015, Dai is appointed a NASA Einstein Fellow at the Institute for Advaned Study, Princeton and is funded to conduct postdoctoral research in cosmology and astrophysics.